Theoretical Characterization of Organic Molecules on Metallic Surfaces: Adsorption, Assembly, Chirality

Dissertation
zur
Erlangung der naturwissenschaftlichen Doktorwürde
(Dr. sc. nat.)
vorgelegt der
Mathematisch-naturwissenschaftlichen Fakultät
der
Universität Zürich

von
Aliaksandr V. Yakutovich
aus
Weißrussland

Promotionskommission

Prof. Dr. Jürg Hutter (Vorsitz)

Prof. Dr. Karl-Heinz Ernst

Dr. Daniele Passerone (Leitung der Dissertation)

Zürich, 2018

Contents

1 Introduction **1**
- 1.1 Model experiments . . . 1
- 1.2 Intermetallic compounds . . . 2
- 1.3 PdGa intermetallic compound . . . 4
- 1.4 Hexabenzocoronene adsorbed on a Pt(111) surface . . . 8
- 1.5 Scanning Probe Microscopies . . . 9
- 1.6 AiiDA . . . 11

2 Methods **13**
- 2.1 Density functional theory: the basics . . . 13
- 2.2 Approximations to exchange-correlation functionals . . . 16
- 2.3 vdW interactions . . . 19
- 2.4 DFT in HPC codes . . . 28
- 2.5 Materials properties from DFT . . . 32
- 2.6 Computation of effective point charges . . . 34
- 2.7 Characterization of molecular transformations at surface . . . 35
- 2.8 Nudged elastic band method . . . 38
- 2.9 Scanning tunneling microscopy simulations . . . 41
- 2.10 Noncontact atomic force microscopy simulations . . . 44

3 Adsorption of 9-EP molecules on PdGa:A(111) surface **49**
- 3.1 Adsorption geometry of 9-EP on a Pd_1 surface . . . 50
- 3.2 Charge density difference in the adsorption of 9-EP . . . 52
- 3.3 $R \leftrightarrow S$ transformation mechanism . . . 54
- 3.4 Origin of the high enantioselectivity . . . 57
- 3.5 Methods . . . 64
- 3.6 Conclusions . . . 67

4 Adsorption of 9-EP molecules on PdGa:A($\bar{1}\bar{1}\bar{1}$) surface **69**
- 4.1 9-EP trimers adsorbed on a Pd_3 surface . . . 70
- 4.2 Monte Carlo study of trimers formation . . . 74
- 4.3 Stabilization of the 9-EP trimer by intermolecular forces . . . 77
- 4.4 Stabilization of the 9-EP trimer by Pd adatom(s) . . . 79
- 4.5 Stabilization of the 9-EP trimer by a Ga adatom . . . 80
- 4.6 Stabilization of the trimer by CO molecule . . . 82
- 4.7 R $\leftrightarrow$ S transformation mechanism . . . 85
- 4.8 Methods . . . 86
- 4.9 Conclusion . . . 87

5 Adsorption of Hexabenzocoronene on Pt(111) surface **89**
5.1 Low HBC coverage on a Pt(111) surface . . . 90
5.2 High HBC coverage on a Pt(111) surface . . . 92
5.3 Origin of the intermolecular repulsion . . . 94
5.4 Methods . . . 99
5.5 Conclusion . . . 100

6 ProbeParticle model: evaluation and improvement **103**
6.1 nc-AFM models . . . 105
6.2 CO polarization and RESP charges . . . 106
6.3 Tip parameters for PP and 2PP models . . . 108
6.4 Case study: PTCDA . . . 109
6.5 Case study: Pentacene . . . 110
6.6 Case study: FFPB . . . 112
6.7 Case study: formic acid dimer . . . 113
6.8 Case study: phtalocyanine . . . 114
6.9 Fitted Lennard-Jones parameters . . . 116
6.10 Methods . . . 117
6.11 Conclusion and outlook . . . 119

7 AiiDA **123**
7.1 ADES model . . . 124
7.2 AiiDA – a Python implementation of the ADES model . . . 125
7.3 AiiDA in the nanotech@surfaces laboratory . . . 128
7.4 CP2K plugin for AiiDA . . . 128
7.5 Workflow to study the electronic properties of 1D materials . . . 131
7.6 Conclusion . . . 132

Conclusion and outlook **137**

List of publications and author contributions **155**

Acknowledgements **157**

Abstract

Intermetallic compounds (IMC) combine different metals with precise stoichiometry into an ordered crystal structure. IMCs have recently attracted increasing interest in heterogeneous catalysis due to their stability and their crystalline surfaces. The main focus of this work lies on exploring the exceptionally high enantioselectivity of chiral (111) and $(\bar{1}\bar{1}\bar{1})$ surfaces of the IMC PdGa, i.e. the ability of the material to selectively adsorb one particular enantiomer (mirror image) of the 9-ethynylphenanthrene (9-EP) molecule. To pin down the origin of the enantioselectivity it was necessary to carefully describe the weak dispersion interactions between adsorbed molecule and the surface layers of the PdGa(111) substrate. The PdGa$(\bar{1}\bar{1}\bar{1})$ surface has a different structure, which induces a different mechanism of enantioselective recognition: 9-EP forms trimers composed of molecules in the same chiral form. Combining scanning tunneling microscopy (STM) and noncontact atomic force microscopy (nc-AFM) experiments with density functional theory (DFT) simulations, it was shown that trimer formation is likely stabilized by Ga adatoms.

The hexabenzocoronone (HBC) molecule is a convenient model for investigating the adsorption of nanographenes on metal surfaces. The adsorption of HBC on Pt(111) was investigated in this work. Combining STM experiments with DFT simulations, two symmetrically equivalent adsorption sites of HBC molecules were identified. Adsorption experiments at high coverage revealed repulsive intermolecular interactions whose origins were investigated using DFT. It was shown that the HBC molecule transfers electron density to the Pt(111) surface. When HBC molecules approach on the surface, the overlap of this densities leads to the Pauli repulsion.

As part of this work, a new efficient simulation protocol for nc-AFM simulations was developed, bridging the gap between accurate, but slow DFT-based simulations and existing classical models. It is shown that the new model provides both qualitative and quantitative improvements over the existing model at comparable computational cost.

The last part of the thesis focuses on the move from individual, manually set up calculations to the automation of recurring simulation tasks using the AiiDA framework. An automated workflow for computing the band structure of graphene nanoribbons has been developed, including an AiiDA plugin for the popular CP2K code.

Kurzfassung

Intermetallische Verbindungen (IMV) sind Kristalle, die verschiedene Metalle in präziser Stöchiometrie enthalten. Die Stabilität der IMV, zusammen mit ihren kristallinen Oberflächen, machen diese Materialien interessant für die heterogene Katalyse. Diese Arbeit beschäftigt sich mit der Enantioselektivität der chiralen (111) and ($\bar{1}\bar{1}\bar{1}$) Oberflächen des IMV PdGa, d.h. mit der Fähigkeit diese Materials selektiv ein Enantiomer (Spiegelbild) des 9-Ethynylphenanthren (9-EP) Moleküls zu adsorbieren. Eine adäquate Beschreibung der Ursachen der Enantioselektivität erforderte eine genaue Beschreibung der van-der-Waals Wechselwirkungen zwischen dem adsorbierten Molekül und den Atomlagen der PdGa (111) Oberfläche. Die PdGa($\bar{1}\bar{1}\bar{1}$), im Gegensatz, zeigt einen anderen Mechanismus für die enantioselektive Adsorption: Hier finden sich jeweils drei 9-EP derselben chiralen Form zu einem Trimer zusammen. Die Erkenntnisse aus Rastertunnelmikroskopie (RTM), Rasterkraftmikroskopie (RKM) und Dichtefunktionaltheorie (DFT) legen nahe, dass die Bildung der Trimere durch Ga Atome stabilisiert wird.

Das Hexabenzocoronene (HBC) Molekül eignet sich als Modell, um die Adsorption von Nanographen-Strukturen auf Metalloberflächen zu charakterisieren. In dieser Arbeit wurde die Adsorption von Hexabenzocoronene (HBC) auf der Platin (111) Oberfläche untersucht. Durch Vergleich von RTM und DFT Rechnungen wurden zwei symmetrisch äquivalente Adsorptionspositionen identifiziert. Weiter weisen Adsorptionsexperimente bei hoher Bedeckung auf abstossende intermolekulare Wechselwirkungen hin. DFT Rechnungen zeigen, dass HBC Elektronendichte an das Substrat abgibt. Nähern sich HBC Moleküle auf der Oberfläche, so fürt der Überlapp dieser induzierten Elektronendichte zur Pauli-Repulsion.

Als Teil dieser Arbeit wurde ein neues Modell für RKM Simulationen entwickelt, welches die Lücke zwischen genauen, aber langsamen DFT-basierten Simulationen und existierenden klassischen Modellen überbrückt. Das Modell liefert sowohl qualitative als auch quantitative Verbesserungen gegenüber bisherigen Modellen, bei vergleichbaren Rechenkosten.

Der letzte Teil der Arbeit geht auf den Weg von einzelnen, manuell aufgesetzten Rechnungen zur Automation wiederkehrender Simulationsschritte mit Hilfe der AiiDA Software. Entwickelt wurden ein automatischer workflow zur Berechnung der Bandstruktur von Graphen Nanostreifen wurde entwickelt, sowie ein AiiDA Plugin für die populäre CP2K software.

Preface

A computational scientist in a mostly experimental facility can help both in the explanation of the observed phenomena and in the prediction and design of novel experiments. This is also the case of my research in a nano- and surface science laboratory at Empa. Across my whole doctoral appointment, I have been striving to validate the adopted levels of theory and methods against experimental results, to develop efficient algorithms for *in silico* analytical tools (e.g., noncontact atomic force microscopy, nc-AFM) and to automate certain computational tasks that are recurrent while interplaying with atomic precision experiments.

The work discussed in the thesis is mostly theoretical, but since it was done in close collaboration with experimentalists, the experimental results are also provided where necessary.

The manuscript contains seven chapters. The first chapter provides an introduction to the problems that are aimed to be solved in this work. The following methods chapter gives an overview of the theoretical tools used throughout this study. The next three chapters discuss the attempts to provide a theoretical explanation of the experimentally observed phenomena such as the enantioselectivity of PdGa(111) and PdGa($\bar{1}\bar{1}\bar{1}$) surfaces, as well as the formation of molecular domains on a Pt(111) surface. Chapter six is focused on improving a methodology for simulating of noncontact atomic force microscopy images. The last chapter discusses a recently proposed paradigm change of modern-day approaches of computational science. It demonstrates how the new approach can be beneficial not only for individual scientists but also for the wider scientific community.

List of abbreviations

ADES	Automation, Data, Environment, Sharing
AFM	Atomic force microscopy
DFT-AFM	DFT based noncontact atomic force microscopy simulations
nc-AFM	Noncontact atomic force microscopy
BO	Born-Oppenheimer approximation
CCSD(T)	Coupled cluster single double (triple)
DFA	Density functional approximation
DFT	Density functional theory
FAD	Formic acid dimer
FFPB	4-(4-(2,3,4,5,6-pentafluorophenylethynyl)-2,3,5,6-tetrafluorophenylethynyl) phenylethynylbenzene
FFT	Fast Fourier transforms
GGA	Generalized gradient approximation
HBC	Hexabenzocoronene
HF	Hartree-Fock
HK	Hohenberg and Kohn
HPC	High-performance computing
IMC	Intemetallic compounds
KS	Kohn-Sham
LEED	Low-energy electron diffraction
LDA	Local density approximation
MBD	Many-body dispersion
MC	Monte Carlo method
MD	Molecular dynamics
MEP	Minimum energy path
PP	ProbeParticle model
2PP	2-point-ProbeParticle model
PTCDA	Perylenetetracarboxylic dianhydride
QHO	Quantum harmonic oscillator
RESP	Restrained electrostatic potential
ROHF	Restricted open-shell Hartree–Fock
RT	Room temperature
STM	Scanning tunneling microscopy
TH	Tersoff-Hamann
TST	Transition state theory
hTST	Harmonic approximation to TST
TZV	Triple zeta valence
UHV	Ultra high vacuum
UUIS	Universally unique identifier
vdW	van der Waals interactions
ZPE	Zero-point energy

Chapter 1

Introduction

This chapter gives an introduction to the topics that are the focus of the current work. It starts with a description of the concept of the **model experiment** – an experimental setup that allows to directly compare the measured properties to the results of the simulations. Further the chapter gives a short overview on the intermetallic compounds from the perspective of the catalytic application. Then the discussion leads to a concrete example of a PdGa intermetallic compound – a promising material in the field of enantioselective heterogeneous catalysis. The applicative part of the introduction is finalized with a discussion on the adsorption of Hexabenzocoronene on a Pt(111) surface.

Further the focus shifts towards methodological aspects of the materials modeling. Following the introduction to the noncontact atomic force microscopy (nc-AFM) two models for the nc-AFM simulations (density functional theory based and a purely classical one) are briefly introduced.

The final section of the chapter introduces the modern views on the computational materials science. It highlights issues that the field is facing nowadays and introduces the proposed approaches to solve them.

1.1 Model experiments

In “real life” one deals with objects that cannot be directly modeled due to their large size. Fortunately, modern experimental techniques, such as scanning probe microscopy, have reached an atomic level of precision, allowing for a direct comparison of experimental findings with the results of atomistic simulations.

We define **model experiment** as an experimental setup with (almost) ideal conditions directly comparable to the results of standard *ab initio* DFT simulations: 0 K temperature, 0 bar pressure (ultra-high vacuum) and well-defined crystal structure of the material under study. Of course, experimentally those conditions are unreachable, however it is possible to get very close to them. In the next paragraph we provide an oversimplified example of the experimental setup needed to perform scanning tunneling microscopy (STM) studies of a molecule adsorbed on a metallic surface.

Very low temperatures (which can be achieved by using liquid helium or liquid nitrogen with boiling temperatures of 4 K and 77 K respectively) are typically needed to remove the kinetic energy from the molecules and to provide a higher resolution in the images. To prevent the molecules from stacking in metastable positions, the cooling process usually follows a suitable equilibration at a higher temperature. Ideally those adsorption positions can be directly compared to the results of the geometry optimization of the molecule on the surface. An ultra high vacuum (UHV) can be obtained with the help of ion pumps and is needed to avoid contamination of the surface by the molecules that are present in the air. The Czochralski method can be used to grow crystals along a particular crystallographic direction. In order to obtain an extremely clean and well-structured surface, it is typically spattered by ionized atoms of a nobel gas (like Ar^+) followed by laser annealing of the surface.

The STM measurements are based on the registration of the tunneling current passing from the extremely sharp metallic tip to the metallic substrate with piezoelectric control of its position. Extreme sharpness of the tip (ideally one atom at the apex) is typically achieved by electrochemical etching.

1.2 Intermetallic compounds

Metallic catalysts play a very important role in the field of industrial organic synthesis. They are used for such reactions as hydrogenation/dehydrogenation, carbon-carbon bond formation, oxidation etc. However, the field of intermetallic compounds has not been represented well in the literature until recently. Figure 1.1 shows that the field of intermetallic catalysis started to grow in the year 2010.

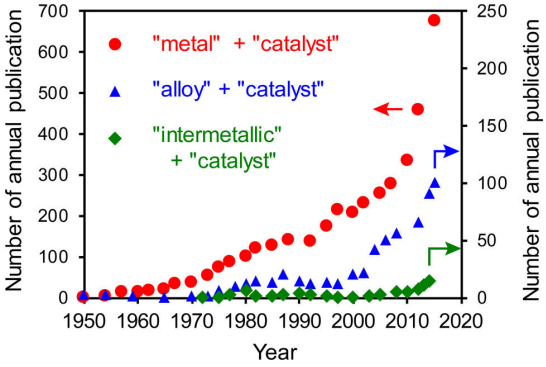


FIGURE 1.1: Annual increase in the number of literature works in which titles include "catalyst" and "metallic", "alloy", or "intermetallic", as determined by Web of Science searches. The figure is taken and adapted from the reference [1].

The reason for such growth is the great potential for intermetallic compounds to be used in the field of heterogeneous catalysis. Dealing with pure metals, one can not have much control on the catalytic properties apart from the physical transformations, such as changing the particles' size. Moving from metals containing only one element to bimolecular systems allows to influence the electronic and geometrical environment of the material's surface. This can be achieved by working with alloys. However, the alloys do not have a well defined structure, which creates a principle difficulty in controlling the arrangement of the atoms on the surface.

At the same time, intermetallic compounds keep a well-defined crystal structure, which makes them very promising candidates for heterogeneous catalysis. Furukawa and Komatsu in their review [1] distinguish four different effects that intermetallic compounds have in comparison to single atom metals, namely: an **electronic effect** (or ligand effect), **geometric effect** (or ensemble effect), **ordering effect** and **steric effect**. The first and the second effects are retained by both alloys and intermetallic compounds, while the third and fourth are the pure properties of the intermetallic compounds.

The nature of the **electronic effect** lies in the modification of the electronic states of the parent metal via its interaction with the states of the guest atom. The **geometric effect** is based on the separation of the more active atoms of the parent metal through interaction with the less active guest atom. This modification creates separated atomic ensembles which have less area for adsorption comparing to the pure

metals. Therefore, the species that require a larger adsorption site are less favorable to be adsorbed. Since both effects do not require any ordering of the atomic structure, they can also be attributed to the alloys.

The **ordering effect** can be attributed only to intermetallic compounds. However, to distinguish this effect one should compare a metallic alloy and intermetallic compound with the same composition. The compounds which can have both ordered and disordered structures are not relevant for this study (Au-Cu, Pt-Fe, Pt-Co, and Pt-Cu).

The **steric effect** is the property of an intermetallic compound which imposes some preferences on the geometry of the adsorbate due to the ordered arrangement of atoms. It mainly depends on the space group of the intermetallic compound as well as on its surface structure. Specific atomic arrangement, for example, can enforce the adsorption of one particular enantiomer, increasing the rate of its chemical transformation. Two examples of a selective reaction were reported by Furukawa et al. for the processes of alkene isomerization [2] and the hydrogenation of nitrostyrene [3] on the Rh-based intermetallic compounds. In the first case the selectivity to *trans*-alkenes was demonstrated by the compounds belonging to the *Pnma* space group, while the $Pm\bar{3}m$ space group compounds had shown a preference to the *cis*-alkenes. In the second case Furukawa et al. studied the selectivity in hydrogenation of nitro and vinyl groups of nitrosterene. Authors had demonstrated that the selectivity of the hydrogenation of nitro group varies from 3% for Rh on SiO_2 to 97% for RhIn on SiO_2. The high selectivity of the latter material is originating from its ordered structure favoring the adsorption of the nitro group with respect to the vinyl group.

1.3 PdGa intermetallic compound

PdGa is an intermetallic compound that emerged recently for its catalytic properties in the purification of feedstock for polyethylene production [4, 5, 6, 7]. Rameshan et al. [8] have also shown that PdGa in contact with oxygen, or when supported by metal-oxide substrates, can be used as a CO_2 selective catalyst in the methanol steam reforming process. Moreover, recently Prinz et al. [9] demonstrated that PdGa can

be potentially used as an enantioselective heterogeneous catalyst due to its chiral crystal structure.

The cubic unit cell of PdGa has eight atoms and belongs to the $P2_13$ space group (see Figure 1.2a, inset). It exhibits two enantiomeric forms, labeled **A** and **B** in the literature [10, 11]. In a recent work [12] the surface terminations of PdGa(111) and PdGa($\bar{1}\bar{1}\bar{1}$) surfaces obtained by the sputter-annealing cycles of a PdGa crystal grown by the Czochralski method [13] were determined. The assignments derived in this work were then confirmed by later experiments and simulations [9, 14, 15]. The two surfaces do not present reconstruction, have equal surface symmetry but exhibit different atomic configurations of the Pd atoms. Referring to the **A** form, the most stable ($\bar{1}\bar{1}\bar{1}$) termination (Pd_3 in the literature [10]) is characterized by one Pd trimer per surface unit cell in the outer atomic layer (Figure 1.2b) while the (111) termination (Pd_1 in the literature [10]) exhibits one isolated Pd site per unit cell (Figure 1.2c).

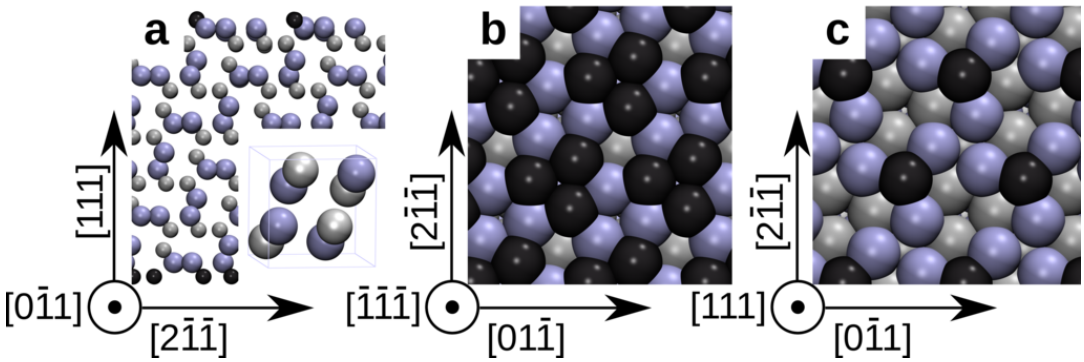


FIGURE 1.2: Atomistic models of a PdGa:A slab and top views of the Pd_3 and Pd_1 surfaces. The gray (black on the surface) and violet spheres represent Pd and Ga atoms respectively. (a) [0$\bar{1}$1] view of a PdGa:A slab. The cell used in our simulations contains 23 atomic layers terminated by the Pd_1 surface on top and the Pd_3 surface on the bottom. The inset shows the PdGa:A bulk unit cell. (b) Top view of the Pd_3 surface with Pd trimers highlighted in black. (c) Top view of the Pd_1 surface with Pd isolated sites highlighted in black.

For both surfaces the second atomic layer consists of Ga trimers, while the third atomic layer is a single Pd atom per unit cell for the Pd_3 surface and a Pd trimer for the Pd_1 surface. Despite the outermost atomic layer of the Pd_1 and Pd_3 surfaces being achiral, the structure of the second and third layer induces a surface chirality. Surface chirality is sketched in Figure 1.3 for the case of the Pd_1 surface. Both **A** and

B forms are presented in the image emphasizing their mirror symmetry. Helixes formed by the Pd and Ga atoms while entering the bulk are highlighted with arrows and different shades (more intense means closer to the reader).

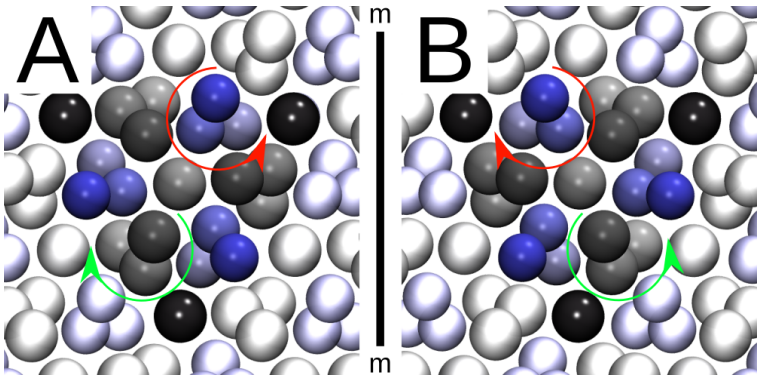


FIGURE 1.3: Top view of PdGa:A(111) and of PdGa:B($\bar{1}\bar{1}\bar{1}$) surfaces (left and right panel respectively). The gray (black at the surface) and violet spheres represent Pd and Ga atoms respectively. The mirror chirality of the two **A** and **B** forms is highlighted by the green and red arrows, following the [$\bar{1}\bar{1}\bar{1}$] direction to enter the bulk, as well as by the intensity of the atoms' color (intensity decays when entering the bulk). The positions of the Pd/Ga atoms describe a clockwise/anticlockwise helix in the case of form **A** and the opposite in the case of form **B**.

The difference in the adsorption energy of a chiral or prochiral molecule on a chiral substrate is at the origin of chiral molecular selection, a process of crucial relevance in diverse chemical fields ranging from pharmaceutics to materials science.

Today it can be optimistically said that the original claim of Jacoby [16], on the lack of exploration of the chiral surface chemistry, was not ignored. Indeed, the theoretical and experimental exploration of chiral inorganic surfaces has been considerably growing in the last decade, particularly with a focus on surface steps (see, e.g., [17]). Most studies, however, exploited the concept that enantioselectivity requires the functionalization of the substrate (through steps, adatoms or, in general, nanopatterning) such that three side groups of the adsorbate molecule are confined within a stereoselective environment [18, 19, 20].

Those concepts should be extended, noting that surface chirality can originate not only from the functionalization of the surface itself (through steps or defects) but also from the natural termination of bulk chirality, as in the intermetallic compound PdGa which is the focus of this work.

Recently, the possibility to exploit the intrinsic bulk chirality of PdGa has been demonstrated by Prinz et al. [9] for the Pd_1 surface. It was proven that adsorption of the prochiral molecule 9-ethynylphenanthrene (9-EP shown in Figure 3.1c) on PdGa:A(111)Pd_1 shows enantiomeric excess (defined as $(n_R - n_S)/(n_R + n_S)$ where n_R and n_S are the number of R and S enantiomers, respectively) of more than 0.9 for the R surface enantiomers (and for the S surface enantiomer on the mirror symmetric equivalent surface PdGa:B($\bar{1}\bar{1}\bar{1}$)Pd_1). In their work, Prinz et al. study the temperature dependence of the enantioselective process and show that the racemic form obtained after deposition of 9-EP at temperatures below 120 K is transformed to an almost enantiopure ensemble at room temperature. The mechanism responsible for this high enantioselectivity has remained elusive up to now.

In chapter 3 we provide an understanding of the observed enantioselectivity. To explain the observed phenomena, we carried out three subsequent analysis steps: (i) we showed that the assignments made by Prinz et al. [9] on the adsorption geometries of the R and S form of 9-EP on a Pd_1 surface are correct; (ii) we demonstrated how a racemic mixture can be converted into enantiomeric excess; (iii) we identified the molecule-substrate interactions responsible for the enantioselectivity assuming that the experimental scanning tunneling microscopy (STM) images were taken at the thermodynamic equilibrium state. Points (ii) and (iii) required challenging calculations both in terms of computational effort and in terms of accuracy of the underlying methodology.

The enantioselectivity of the PdGa($\bar{1}\bar{1}\bar{1}$)Pd_3 surface is also discussed in the thesis (chapter 4). Interestingly, the mechanism for the enanitoselective recognition turned out to be completely different from the one at Pd_1. As it was shown experimentally, at room temperature the Pd_3 surface does not possess any enantioselectivty. However, at elevated temperatures (> 400 K) 9-EP molecules start to form trimers almost exclusively formed by the molecules in R configuration. The very first question that arose from that point was: what happens in the center of the 9-EP trimer? STM and noncontact atomic force microscopy (nc-AFM) experiments could not clearly validate any of the proposed hypotheses (even though nc-AFM experiments were

particularly helpful in excluding the assumption of alcohol formation). We combined theoretical and experimental approaches in order to get an idea about the exact chemical and structural nature of the trimer center. According to the thermodynamical estimation, the driving force for trimer formation can not be a simple vdW interaction among the molecules. Therefore, we had to guess a chemical transformation that may lead to such a change. We tested several possible structures and came to the conclusion that the only real possibility may be the stabilization of the 9-EP trimer by Ga adatom.

1.4 Hexabenzocoronene adsorbed on a Pt(111) surface

An understanding of the self-organization of molecules on low-index metallic surfaces is crucial for fields of high technological importance such as organic photovoltaics, molecular switches, field effect transistors, etc. In this work the adsorption of Hexa-peri-hexabenzocoronene (HBC) on a Pd(111) surface was explored. In addition to the possible technological applications the molecule is a convenient model for investigating the adsorption of nanographenes on metal surfaces.

With the help of STM, two different configurations of the HBC molecule were detected on a bare Pt(111) surface at low coverage. To clearly identify those configurations several different initial guesses of the adsorption geometry were optimized using DFT. As a result it was found that two symmetry-related (therefore equivalent energetically) geometries appeared to be the most stable in a good agreement with the experimental observations.

The second question to answer was to understand the repulsive nature of the intermolecular interactions. The latter was assumed based on the experimental evidence: no clustering of 9-EP at low coverage, and an absence of close-packed structure at high (almost 1 monolayer) coverage, was observed. To understand this behavior the DFT energies of the HBC dimers at different adsorption positions were computed. It was indeed found that for the close-packed structures the energy was considerably higher compared to the less dense ones. The charge-density difference technique allowed to identify an electron density transfer from the HBC molecule to

the (111) surface of Pt. The overlap of the transferred charge-density caused an increase of the Pauli repulsion, explaining the instability of the close-packed structures of HBC.

The detailed description of this work can be found in the chapter 5

1.5 Scanning Probe Microscopies

The era of scanning probe microscopies started with the invention of the first scanning tunneling microscope (STM) at IBM Rüschlikon by Binnig and Rohrer in the year 1983 [21]. The first patent disclosure for the STM microscope, however, had been submitted by the authors about 4 years before, in January 1979. This revolutionary technique allowed to achieve unprecedented resolution and provided access to the atomic structure of materials.

In 1986 Binnig et al. [22] proposed a modification of the combination of the scanning tunneling microscope with the stylus profilometer. This new approach allowed to measure atomic forces as small piconewtons [23][1]. The new technique was named atomic force microscopy (AFM). Contrary to STM, where a bias voltage is applied to the substrate/tip system and a tunneling current is measured, in AFM the force acting on the tip is measured in absence of any additional influence. The technique was further improved by Martin et al. [24], who introduced a vibrating tip and proposed a simple theoretical model that connects a change in vibrational frequency ($\Delta\omega$) to a variation of the derivative of the vertical force acting from the substrate to the tip. The application of noncontact dynamic force microscopy has been further expanded to study biological systems such as DNA and the tobacco mosaic virus [25].

Nevertheless, the ultimate goal of AFM was set to achieve atomic resolution, which was a challenging task due to the complex nature of the forces between the tip and sample as well as technical difficulties to measure those forces with extreme precision. One of the first successful atomic resolution images were reported by Giessibl [26], who could visualize the reactive surface of Si(111) by AFM operated in noncontact mode (nc-AFM). A further significant achievement was made by Sugimoto et al. [27] who in 2007 were able to recognize the chemical identity of surface

[1] While Binnig et al. [22] originally reported the value of 10^{-18} N.

atoms. It was demonstrated that short-range forces between the tip and an atom on the surface (the force arising at the onset of a chemical bond formation between tip and surface atom) depend strongly on the chemical nature of the atoms. The authors outlined a calibration (normalization) procedure that allowed to remove the dependence of the experimental result from the specific setup of the instrument. One year later Sugimoto et al. [28] showed the possibility of atomic manipulations by vertical interchange of atoms between the surface of a semiconductor and the apex of an AFM tip.

However, the application of nc-AFM microscopy for imaging of small organic molecules adsorbed on the surface was impossible due to the high chemical reactivity of the metallic tip. This problem was resolved later by Gross et al. [29] who functionalized the nc-AFM tip with a CO molecule. The molecule confers to the tip a well-defined sharp geometry. The inertness of the CO allows to the tip to approach the surface at close distances, granting access to structural imaging with unprecedented atomistic details. Further refinements of the technique allowed to discriminate for "bond order" in chemical bonds [30], to image the charge distribution within a molecule [31], and to map the electrostatic field of molecules [32]. A standard setup for nc-AFM with CO functionalization of the tip is shown in Figure 1.4.

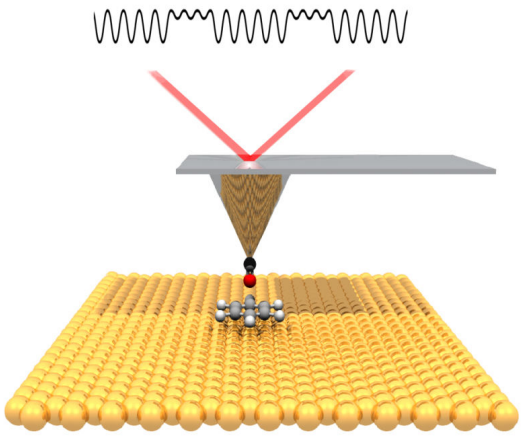


FIGURE 1.4: A schematic representation of an experimental nc-AFM setup. The cantilever with a CO molecule at the apex of the metallic tip acts as a probe to scan the surface of the planar molecule adsorbed on the material. *Credit: Lawrence Berkeley National Laboratory.*

The cantilever is set to oscillate with a proper frequency ω_0, and the force experienced by the tip due to the CO/substrate interaction induces a frequency shift that

is recorded while scanning the sample.

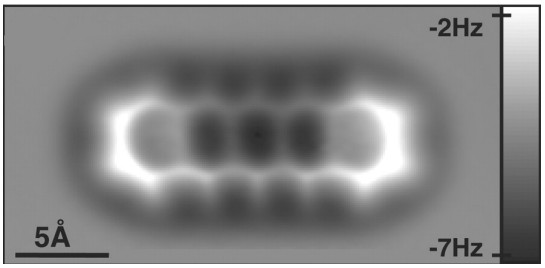


FIGURE 1.5: First high-resolution nc-AFM image of Pentacene molecule obtained by Gross et al. [29]. Figure was taken and adapted from the reference [29].

In this manuscript, our attempts to improve a classical model for nc-AFM simulations are discussed. The work started with a comparison of the classical so-called ProbeParticle (PP) model [33], which has been successfully used in different studies [34, 32]. The model was pushed towards its limits by a fitting of its parameters to the results of DFT-based AFM simulations. The latter was used as a "gold standard" in both a qualitative and quantitative respects. It was first demonstrated that the standard PP model had some limitations in predicting the features of nc-AFM imaging. To overcome them it was proposed to enhance the PP model via introduction of the second probe particle in the tip, representing the carbon atom. This modification has resulted in a better agreement with the DFT and ability to better reproduce the sharp features observed experimentally.

The details of the work are discussed in the chapter 6.

1.6 AiiDA

The development of simulation techniques in computational materials science together with the increase of computational power available to researchers has led, in the last decade, to a dramatic increase of data describing materials properties. Parallel to the continuous improvement of the accuracy in simulations, new challenges appear in (i) handling the huge amount of data available (ii) sharing it among researchers (iii) checking the accuracy of data in large databases (iv) reproducing results that are obtained from a complex chain of calculations, each one depending on the other. We are disposing almost unlimited data storage capabilities, however we

are still lacking the ability to manage, share and analyze the stored data. Therefore, the current approach to computational research needs to be "redesigned".

Different initiatives have been launched to overcome those problems by (i) creating repositories for comprehensive storage, management and sharing of scientific data [35, 36, 37, 38, 39, 40] (ii) developing a general format to store the results computed with commonly used quantum chemistry packages [41] (iii) constructing an infrastructure for high-throughput [42, 43] and distributed [44] calculations which become feasible due to the exponential increase of the computational power. In a recent project named AiiDA [45] (Automated Interactive Infrastructure and Database for Computational Science) the three issues above are tackled together: "AiiDA is a flexible and scalable informatics' infrastructure to manage, preserve, and disseminate the simulations, data, and workflows of modern-day computational science ..." [46].

In chapter 7 the advantages of using the AiiDA package are discussed. The chapter starts with a description of the ADES [2] model and a brief introduction on how automation, data management, a common environment for different codes and sharing of the results are adressed within this project. Then, providing an example of CP2K integration, it will be shown how a DFT-code can be embedded in the AiiDA framework. Finally, the description of a workflow that was designed and developed to fit the needs of the nanotech@surfaces laboratory at Empa will be provided.

[2] Automation, Data, Environment, Sharing.

Chapter 2

Methods

This chapter discusses the theoretical details of the simulation methods, mostly relying on density functional theory [47, 48] (DFT), that have been used in current research work. After an introduction to the DFT approximations, the focus will be shifted to a description of simulation approaches that allow to compare simulations and experiments in terms of thermodynamic properties and microscopy features. In the description of DFT, particular attention will be dedicated to the discussion of van der Waals interactions (vdW) that contribute substantially in determining the geometrical (correspondingly electronic) properties of molecule/substrate interfaces.

2.1 Density functional theory: the basics

When exploring the properties of systems ranging from simple molecules to multicomponent bulk systems, approximations to density functional theory are considered as the standard theoretical approach for computational materials science and chemistry. To understand the hypotheses underlying DFT it is useful to consider a material (or a molecule) as a system of N electrons that are interacting under the influence of an external potential, originating from M atomic nuclei. If the nuclei

are considered as classical point charges whose motion is well described by classical equations of motion where the forces on the nuclei originate from the electrostatic nuclei-nuclei interactions and the electron-nuclei interactions[1], then the Born-Oppenheimer approximation [51] allows to write the Hamiltonian for the electrons of the system as

$$\hat{H} = \underbrace{-\frac{1}{2}\sum_{i=1}^{N}\nabla_i^2}_{\hat{T}} + \underbrace{\sum_{i>j}^{N}\frac{1}{|\vec{r}_{ij}|}}_{\hat{U}} - \underbrace{\sum_{i,\alpha}^{N,M}\frac{Z_\alpha}{|\vec{r}_{i\alpha}|}}_{\hat{V}} \tag{2.1}$$

where $\hat{T}$ is the kinetic energy operator, $\hat{U}$ is the electron-electron interaction operator, $\hat{V}$ is the electron-ion interaction (external potential) operator, $|\vec{r}_{ij}|$ is the distance between electron i and electron j, Z_a is the atomic number of nucleus α, $|\vec{r}_{i\alpha}|$ is the distance between electron i and nucleus α. In 1964 Hohenberg and Kohn (HK) demonstrated that the ground state properties of a system of interacting electrons affected by an external potential derive from a unique functional of the electron density $n(\vec{r})$, where [2]

$$n(\vec{r}) = N\int \mathrm{d}^3 r_2 \cdots \int \mathrm{d}^3 r_N \Psi^*(\vec{r},\vec{r}_2,\ldots,\vec{r}_N)\Psi(\vec{r},\vec{r}_2,\ldots,\vec{r}_N)^3 \tag{2.2}$$

with $\Psi(\vec{r_1},\vec{r_2}\ldots,\vec{r_N})$ the wave function of the system of N electrons. Moreover, they demonstrated that a functional of the electron density of the system exists such that its minimum value, corresponding to the ground state energy of the system, is obtained at the ground state charge density. The HK theorem is the basis of DFT. The charge density functional can be written as

$$E[n] = T[n] + U[n] + V[n]^4 \tag{2.3}$$

where

$$V = \int v(\vec{r})n(\vec{r})d\vec{r} \tag{2.4}$$

with $v(\vec{r})$ taking the form of $\sum_\alpha^M \frac{-Z_\alpha}{|\vec{R}_\alpha - \vec{r}|}$ in the case of nuclei and electrons not influenced by an external field. The total energy of the system is conveniently rewritten

[1] The nature of the electron-nuclei interaction will be discussed in more detail in section 2.5, where the Hellman-Feynman [49, 50] theorem is introduced.

[2] The spin coordinates of the electrons are neglected for simplicity.

[3] This equation is correct for a normalized ψ.

[4] Even though the form of the equation is similar to the one of Hamiltonian, there is a significant difference: now it contains functionals of the charge density but not operators.

as

$$E = V[n] + F[n] \tag{2.5}$$

where $F[n] = \langle \Psi | \hat{T} + \hat{U} | \Psi \rangle$. Writing $F[n]$ in the form

$$F[n] = \int \frac{n(\vec{r})n(\vec{r'})}{|\vec{r} - \vec{r'}|} d\vec{r'} + T_0[n] + E_{xc}[n] \tag{2.6}$$

allows to separate the so-called exchange-correlation functional ($E_{xc}[n]$), which is unknown, from the classical Coulomb energy and the kinetic energy ($T_0[n]$) of a non-interacting electron gas of density $n(\vec{r})$. Before DFT can be effectively considered as a route to access the ground state properties of materials, a way to minimize the energy functional and to represent the unknown exchange correlation part has to be defined. The minimization procedure has to be transferred to efficient computational schemes. One year after the HK theorem, Kohn and Sham [48] introduced a single-particle scheme, within the DFT formalism, that allows to represent the ground state density of a system in terms of an auxiliary system of "non-interacting electrons" with single particle wave functions $\psi_i(\vec{r})$ and to solve self-consistently a set of single-particle-like equations to find the minimum of the HK functional. The charge density of the auxiliary system is given by

$$n[\vec{r}] = \sum_i^n |\psi_i(\vec{r})|^2 \tag{2.7}$$

$\psi_i(\vec{r})$, called Kohn-Sham (KS) orbitals, are the N (number of electrons in the system) solutions of the single particle equations

$$\left[-\frac{1}{2}\nabla^2 + [\phi(\vec{r}) + \mu_{xc}(n)] \right] \psi_i(\vec{r}) = \epsilon_i \psi_i(\vec{r}) \tag{2.8}$$

where $\phi(\vec{r}) = v(\vec{r}) + \int \frac{n(\vec{r'})}{|\vec{r} - \vec{r'}|} d\vec{r'}$. Equation 2.8 is only apparently a system of independent equations, the mathematical complexity of the problem is condensed in the $\mu_{xc}(n)$ term defined as

$$\mu_{xc}(n) = \frac{\delta E_{xc}[n]}{\delta n(\vec{r})} \tag{2.9}$$

μ_{xc} depends on $n(\vec{r})$ and, as such, on all the single particle wave functions, thus

the system of equations has to be solved iteratively. The complexity of the electron-electron interaction is hidden inside the so-called exchange-correlation potential, which does not have a well-defined mathematical form. Therefore, various approximations to DFT are present in the literature. Some of them are briefly discussed in the following section.

2.2 Approximations to exchange-correlation functionals

The exchange correlation functional has been (and still is) the subject of several research works attempting to define approximations more and more accurately and, possibly, not specific to a restricted class of materials.

The first functional adopted by the DFT community was derived from the so-called local density approximation (LDA) where the electron gas is considered as locally homogeneous and the exchange-correlation energy is obtained as [48]

$$E_{xc}^{LDA}[n(\vec{r})] = \int \varepsilon_{xc}(n) n(\vec{r}) d\vec{r} \tag{2.10}$$

where the right side of the equation is the first term in the expansion of the $E_{xc}[n]$ functional

$$E_{xc}[n] = \int \varepsilon_{xc}(n) n(\vec{r}) d\vec{r} + \int \varepsilon_{xc}^{(2)}(n) |\nabla n(\vec{r})|^2 d\vec{r} + \dots \tag{2.11}$$

In the Equation 2.10 the non-local contributions to the functional are neglected, making it dependent only on the charge-density value at every point in space. Naturally, the non-locality can be introduced back making the functional dependent on $\nabla n(\vec{r})$ and higher order derivatives of $n(\vec{r})$. This gives rise to more advanced DFT approaches such as generalized gradient approximation (GGA) and others, which are discussed further in the section in more detail.

Therefore, instead of describing the exchange and correlation of inhomogeneous electron density, at each point the exchange-correlation energy of a corresponding homogeneous system with the *real* density at that point is considered. For magnetic systems, where spin polarization has to be taken into account, the local spin density approximation (LSDA) is introduced and the exchange-correlation energy E_{xc}

is a functional of the local electron-spin densities $n_{\uparrow}(\vec{r})$ and $n{\downarrow}(\vec{r})$. To avoid making the notation too heavy, the spin degrees of freedom will not be considered in this paragraph. $E_{xc}[n]$ can be written as the sum of the exchange and correlation contributions

$$E_{xc}[n] = E_x[n] + E_c[n] \tag{2.12}$$

Knowing the density of the homogeneous electron gas as

$$n(\vec{r}) = \frac{1}{\frac{4}{3}\pi r_s^3} \tag{2.13}$$

where r_s is the Wigner-Seitz radius of the sphere occupied by each electron, the exchange energy of a homogeneous gas was written as

$$E_x = \frac{0.4582}{r_s} \tag{2.14}$$

Please note that the exchange energy can now be substituted into the integral 2.10 with a dependence on the density $n^{\frac{1}{3}}$.

A very simple formula for the correlation was derived by Chachiyo [52]

$$E_c = a \ln\left(1 + \frac{b}{r_s} + \frac{b}{r_s^2}\right) \tag{2.15}$$

with $a = \frac{\ln(2)-1}{2\pi^2}$ and $b = 27.420$ (in atomic units). This correlation functional agrees very well with the result of quantum Monte-Carlo simulation by Ceperley and Alder [53]. The LDA approximation is accurate for systems with a slow varying charge density. The assumption of homogeneity can however be too hasty in many cases, for example in molecules. A step towards more accurate functionals is to include the gradient of the density in $E_{xc}[n]$

$$E_{\text{xc}}^{GGA}[n] = \int \varepsilon_{xc}(n, \vec{\nabla} n) n(\vec{r})\, d\vec{r} \tag{2.16}$$

This approach, known as generalized gradient approximation (GGA) allows to partially account for the non-local nature of the exchange-correlation functional and, in general, provides better results in comparison to the LDA. Within the family of GGA functionals, the one proposed by Perdew, Burke, Ernzerhof (PBE) is one of the most

commonly adopted

$$E_x^{PBE} = 1 + \kappa - \frac{\kappa}{1 + \frac{\mu s^2}{\kappa}} \tag{2.17}$$

with $\kappa = 0.804$ and $\mu = 0.220$. For the exact form of the correlation term of the PBE functional, the reader is referred to the original article [54].

Both previous approximations provide a local description of the exchange operator and fail to cancel the electron self interaction introduced with the classical Coulomb integral. To account for these problems, which can be crucial in systems such as molecules or defected bulks where charge localization is pronounced, starting with the work of Axel Becke [55], a successful strategy was introduced to adopt an "exact"[5] exchange formulation based on the Hartree-Fock formalism

$$E_x^{HF} = -\frac{1}{2}\sum_{i,j}\int \psi_i^*(\vec{r})\psi_j^*(\vec{r})\frac{1}{r_{12}}\psi_i(\vec{r}')\psi_j(\vec{r}')d\vec{r}d\vec{r}' \tag{2.18}$$

where $\psi(\vec{r})$ are the one-electron KS orbitals. The general form of functionals obtained with this recipe, so-called *hybrid* functionals, depends on a weighted sum of a GGA exchange term and the corresponding HF exchange term

$$E_x^{hybrid} = aE_x^{HF} + bE_x^{GGA} \tag{2.19}$$

The two most common examples of hybrid functionals are B3LYP and PBE0. The first was proposed by Stephens et al. [56] and the parameters in its expression

$$E_{xc}^{B3LYP} = E_x^{LDA} + a_0(E_x^{HF} - E_x^{LDA}) + a_x(E_x^{GGA} - E_x^{LDA}) + E_c^{LDA} + a_c(E_c^{GGA} - E_c^{LDA}) \tag{2.20}$$

were fitted to experimental data, conferring to the functional the appellative of "semi-empirical". The PBE0 functional was introduced by Adamo and Barone [57] and by Ernzerhof and Scuseria [58] with the following expression:

$$E_{xc}^{PBE0} = E_{xc}^{PBE} + \frac{1}{4}\left(E_x^{HF} - E_x^{PBE}\right) \tag{2.21}$$

[5] Still the correction is based on the one-electron KS orbitals.

At this point, the numerical complexity of the algorithms behind different functionals should be mentioned. For a system containing N electrons, a straightforward implementation of the KS-DFT would result in $\mathcal{O}(N^3)$ scaling[6], while the HF method would scale as $\mathcal{O}(N^4)$[7] (and so do the hybrid functionals). Linear-scaling algorithms for the DFT and HF methods are also available, however they were not employed in this work since they become an algorithm of choice only for systems more then tens of thousands of atoms. For the interested reader, more details about linear-scaling DFT and HF algorithms and their applicability can be found in the following works: [59, 60, 61].

Crucial to the work discussed in this thesis is the following aspect: the non-local nature of the electron correlation. LDA, GGA functionals are inherently local so they do not account for the correlation. Hybrid functionals only introduce non-local exchange, but do not provide any improvements for the correlation. Thus, the dispersion part of the vdW interactions, originating from the correlation, is not reproduced in the DFT formalism [62] discussed so far. As for the description of molecule-substrate interfaces (studied in this thesis), the dispersion forces play a fundamental role. The next section is fully devoted to the approaches of computing them within the DFT framework.

2.3 vdW interactions

In many problems related to solid state physics or molecular chemistry, neglecting van der Waals interactions does not modify the general scenario. In other cases such as protein folding, packing of molecular crystals or physisorption phenomena, vdW can have a relevant impact. It is known for more then two decades that traditionally used DFT functionals fail in characterizing the long-range dispersion interactions [63]. The ideally expected decay depends on the details of the systems as evident considering extremal cases such as for 2D-2D systems (as for example two sheets of graphene) where the interaction should decay as d^{-4} with the planar

[6] Since the orthogonalization of Kohn-Sham orbitals would require to compute $\mathcal{O}(N^2)$ scalar products, while the computation of a scalar product itself has the complexity of $\mathcal{O}(N)$.

[7] Due to computation of four-center two-electron integrals.

separation d, or for 1D-1D systems (for example carbon nanotubes) where the interaction should decay as l^{-5} with the distance l between the axes of the two tubes or for 0D-0D systems (like in the case of two fullerenes) where the decay should be r^{-6} with the intermolecular distance r [64]. In the case of molecules on surfaces of bulk systems the decay is expected to be d^{-3}. Please note, that for certain systems the real decay can deviate significantly from the ideal one, as it was reported recently by Gobre and Tkatchenko [65].

The divergence of standard DFT, in describing the energy of separation for two systems interacting via dispersion forces, from the correct asymptotic behavior is severe and dominated by an exponentially decaying character. Two examples of such deviation are shown in Figure 2.1: the potential energy curve for two atoms of Kr and potential energy curve for the benzene dimer. In both cases, the nature of the intermolecular interaction is purely dispersive and the curve should have a minimum at a certain distance.

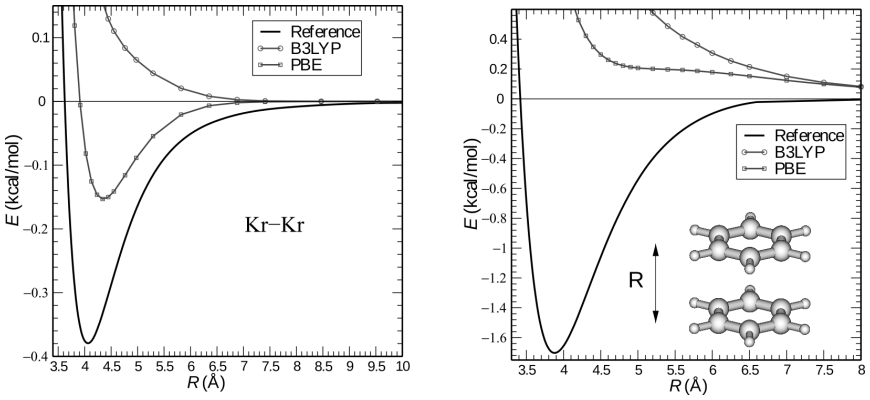


FIGURE 2.1: Potential energy curves for two atoms of Kr (left) and the benzene dimer (right) with two different density functional approximations in comparison with accurate CCSD(T) reference data. The figure is taken from the reference [66].

The origin of the problem lies in the inability of DFT to describe instantaneous electron correlations originated from the "virtual" excitations of electrons to the available energy levels [66]. Those excitations happen due to the zero-point energy fluctuations and result in the corresponding change of electron density. Such

correlated fluctuations of the density allow for electromagnetic interactions (or **dispersion interactions**) between atoms and molecules[8]. The described mechanism of dispersion interactions makes evident the fact that *ab initio* theory, able to properly account for the dispersion interactions, must take virtual states into consideration. The commonly used KS-DFT is based on a single Slater determinant, therefore by construction it only considers the occupied states. Thus, to take dispersion interactions into account DFT must be connected with either a parametrized empirical model or with a reference system where such excitations are taken into account (see the discussion about many-body perturbation theory further in the text). The commonly-accepted approaches to introduce dispersion interactions to DFT will be discussed below in more detail.

DFT-D family of methods

To include dispersion interactions in DFT different empirical and semi-empirical approaches have been proposed in recent years. A common practice is to consider atom centered contributions in the form of C_6/R^6 with a cutoff function ensuring that the attraction is quenched in the "close distance region". This approach is computationally efficient and its cost is negligible with respect to the cost of the DFT calculation. A pioneering attempt in this respect to provide a set of parameters to be employed for the majority of elements in the periodic table was provided by Grimme with the so-called DFT-D family of vdW corrections [67, 68, 69] widely adopted in the literature and also in the present work. The DFT-D correction has the form

$$E_{DFT-D} = E_{DFT} + E_{vdW} \tag{2.22}$$

with E_{vdW} defined as

$$E_{vdW} = -s_6 \sum_{i=1}^{N_{at}-1} \sum_{j=i+1}^{N_{at}} f_{damp}(R_{ij}) \frac{C_6^{ij}}{R_{ij}^6} \tag{2.23}$$

[8]Interestingly enough, such correlated fluctuations can exist not only between separated atoms or molecules, but between parts of the same molecule, providing the origin of the intramolecular dispersion effects [66].

where C_6^{ij} is the dispersion coefficient, R_{ij} is the distance between atoms i and j , s_6 is a global scaling factor which depends on the DFT functional used and $f_{damp}(R_{ij})$ is the quenching function [70]

$$f_{damp}(R_{ij}) = \frac{1}{1 + e^{-d/(R_{ij}/R_r - 1)}} \tag{2.24}$$

where R_r is the sum of atomic vdW radii of elements.

For the DFT-D1 [67] method the C_6 coefficients of the elements were taken from the work of Wu and Yang[9] [70], who fitted those parameters to molecular C_6 parameters obtained experimentally [71]. For the DFT-D2 scheme [68] the atomic coefficients were obtained using the following equation

$$C_6^a = 0.05 N I_p^a \alpha^a \tag{2.25}$$

which was derived from the London dispersion formula. I_p^a and α^a are the atomic ionization potential and static dipole polarizability of atom a both computed with PBE0 functional [57, 58], and N has values 2, 10, 18, 36, and 54 for the first five rows of the periodic table. The vdW radii of all the elements was obtained as a rescaled radius of 0.01 a.u. electron density contour computed with ROHF/TZV[10]. A scaling factor of 1.22 was chosen for the DFT-D1 method and 1.10 for DFT-D2.

The later improved correction (DFT-D3) proposed by Grimme [69] introduced several changes: (i) atom pairwise specific[11] dispersion coefficients and cutoff radii computed from first principles; (ii) To account for the geometrical information, the C_6^{ab} term has now become dependant on the coordination number; (iii) three-body terms. The correction now comprises of two-body and three-body terms

$$E_{disp} = E^{(2)} + E^{(3)} \tag{2.26}$$

[9] However, the coefficients were averaged for the different hybridization states of atoms.

[10] Restricted open-shell Hartree–Fock calculations using the valence triple-zeta basis set.

[11] Please note here that for the DFT-D1 and DFT-D2 methods, those C_6 and R were computed for each element separately. To obtain pairwise C_6^{ab} and R^{ab}, several combination rules were applied [67, 68]. In the DFT-D3 method Grimme et al. [69] C_6^{ab} and R^{ab} were computed specifically for each a/b atomic pair.

where the two-body contribution takes the general form

$$E^{(2)} = -\sum_{AB}\sum_{n=6,8,10...} s_n \frac{C_n^{AB}}{R_{AB}^n} f_{damp}(R_{AB}) \tag{2.27}$$

The C_6^{AB} coefficients were computed *ab initio* using the Casimir-Polder formula [72] which connects the polarizabilities ($\alpha^A(i\omega)$ and $\alpha^B(i\omega)$) of two atoms A and B to their dispersion attraction[12]

$$C_6^{AB} = \frac{3}{\pi}\int_0^\infty \alpha^A(i\omega)\alpha^B(i\omega)d\omega \tag{2.28}$$

The higher-order C_n^{AB} parameters, used to adapt potential to a particular form of DFT functional, were obtained through recursive relations. For more details the reader is referenced to the original article [69].

Please note that the polarizabilities can be computed based on the relationships involving unoccupied states, that can be obtained, for example, employing time-dependent DFT [73]. This brings the discussion back to the initial argument that one needs to involve virtual orbitals for computing dispersion interactions.

The three-body term was also considered by Grimme et al. [69]. The authors used the following Axilrod–Teller–Muto formula to access the three-body contribution to the dispersive energy

$$E^{(3)} = \frac{C_9^{abc}(3\cos(\theta_a)\cos(\theta_b)\cos(\theta_c)+1)}{(r_{ab}r_{bc}r_{ca})^3} \tag{2.29}$$

where θ_a, θ_b and θ_c are the angles of a triangle formed by three particles: a, b and c with interatomic distances r_{ab}, r_{bc} and r_{ca}. The C_9^{abc} term was approximated by the geometrical average of three pairwise C_6 coefficients

$$C_9^{abc} = -\sqrt{C_6^{ab}C_6^{bc}C_6^{ca}} \tag{2.30}$$

Nevertheless, it was not recommended to include this $E^{(3)}$ term by default, since for the small and medium size systems its energetic contribution is negligible. But

[12] In the equation 2.28, instead of using the atomic polarizabilities Grimme et al. used the polarizabilities of hydrides removing the contribution from hydrogen atoms. It was done because the polarizabilities computed for isolated atoms are influenced by energetically low-lying states leading to large dispersion coefficients.

more importantly, it is still not much known how this contribution is treated by the standard density functional in the overlapping region.

Tkatchenko-Scheffler and Many-Body Dispersion families of methods

In 2009 Tkatchenko and Scheffler [74] proposed a similar approach where the C_6 coefficients are derived from free-atom parameters and rescaled according to their Hirshfield partitioning $w_A(\vec{r})$ to make them system dependent

$$w_A(\vec{r}) = \frac{n_A^{free}(\vec{r})}{\sum_B n_B^{free}(\vec{r})} \tag{2.31}$$

where $n_{A/B}^{free}(\vec{r})$ are the charge density of a free atom of type A/B in a molecule.

Approximating the frequency-dependent polarizability $\alpha(i\omega)$ by the leading addend in the Padé series and employing the Casimir-Polder integral (as shown in Equation 2.28), Tkatchenko and Scheffler obtained

$$C_6^{AB} = \frac{3}{2}\left[\frac{\eta_A \eta_B}{\eta_A + \eta_B}\right] \alpha_A^0 \alpha_B^0 \tag{2.32}$$

which relates the C_6^{AB} with the static polarizability $\alpha_{A/B}^0$ and an effective frequency $\eta_{A/B}$. For $A = B$ Equation 2.32 becomes

$$C_6^{AA} = \frac{3}{4}\eta_A (\alpha_A^0)^2 \Longleftrightarrow \eta_A = \frac{4}{3}\frac{C_6^{AA}}{(\alpha_A^0)^2} \tag{2.33}$$

Replacing $\eta_{A/B}$ in Equation 2.32 with Equation 2.33 allows to obtain

$$C_6^{AB} = \frac{2 C_6^{AA} C_6^{AA}}{\frac{\alpha_B^0}{\alpha_A^0} C_6^{AA} + \frac{\alpha_A^0}{\alpha_B^0} C_6^{BB}} \tag{2.34}$$

The free atomic $\alpha_{A/B}^0$ and $C_6^{AA/BB}$ parameters in Equation 2.34 were obtained from the database of Chu and Dalarno [75] for isolated atoms. The approach is extended to molecules and solids employing a direct relation between polarizabilities and effective volumes of atoms inside a molecule

$$\frac{\kappa_A^{mol}}{\kappa_A^{free}} \frac{\alpha_A^{mol}}{\alpha_A^{free}} = \frac{V_A^{mol}}{V_A^{free}} = \frac{\int r^3 w_A(\vec{r}) n(\vec{r}) d^3\vec{r}}{\int r^3 n_A^{free}(\vec{r}) d^3\vec{r}} \tag{2.35}$$

where $\kappa_{A/B}$ is the proportionality constant between atomic volume and polarizability of atom A/B, $n(\vec{r})$ is the electron density of the molecule, $n_A^{free}(\vec{r})$ is the electron density of the free atom. To extract the effective atomic volume the authors employed the Hirshfield atomic partitioning weight defined in Equation 2.31.

Knowing how C_6^{AA} is defined as a function of η_A and α_A^0 (Equation 2.33), it is possible to express the C_6 coefficient for an atom in the molecule ($C_{6,mol}^{AA}$) in terms of the C_6 coefficient of isolated atoms ($C_{6,free}^{AA}$)

$$C_{6,mol}^{AA} = \frac{\eta_A^{mol}}{\eta_A^{free}} \left(\frac{\kappa_A^{free}}{\kappa_A^{mol}} \right)^2 \left(\frac{V_A^{mol}}{V_A^{free}} \right)^2 C_{6,free}^{AA} \tag{2.36}$$

where $\frac{\eta_A^{mol}}{\eta_A^{free}} \left(\frac{\kappa_A^{free}}{\kappa_A^{mol}} \right)^2$ constant is set to 1[13].

The damping function used in the TS approach has the form of Equation 2.24.

The vdW radius of the elements in a molecule is obtained through a rescaling of the vdW radius of the free elements defined as half of the distance between two atoms where Pauli repulsion is compensated by the dispersion attraction

$$R_{mol}^0 = \left(\frac{V^{mol}}{V^{free}} \right)^{1/3} R_{free}^0 \tag{2.37}$$

It is worth noting that Tkatchenko and Scheffler pointed out a possible weakness of their approach for metallic systems, where nonaddictive higher-order effects may be relevant [74].

In a later work [76], where the TS approach [74] is combined to the self-consistent screening equation of classical electrodynamics, many-body contributions to the vdW interactions are considered to account for " ...long-range electrostatic screening extending beyond the range of the exponentially decaying atomic densities" .

This model is known as the "many-body dispersion model" (MBD) [76, 77, 78, 79, 80, 81] and can be schematized as follows: First of all, the charge-density $n(\vec{r})$ of the system is partitioned into atomic Hirshfield volumes, which allow to obtain the static dipole atomic polarizabilites (following the TS approach)

[13]The authors show that this choice is appropriate for a set of 1225 molecules. The largest deviation (44%) is reported for H_2 , and for larger molecules starting from CO_2 and N_2 the error is lower than 10%.

$$\alpha_A^{0,mol} = \frac{V_A^{mol}}{V_A^{free}} \alpha_A^{free} = \left(\frac{\int r^3 w_A(\vec{r}) n(\vec{r}) d^3\vec{r}}{\int r^3 n_A^{free}(\vec{r}) d^3\vec{r}} \right) \alpha_A^{0,free} \tag{2.38}$$

For a system with a finite band gap, the collective charge-density fluctuations originate from correlated dipolar excitations [80]. The MBD method parametrizes those excitations by a set of atomically centered quantum harmonic oscillators (QHOs) characterized by frequency-dependent dipole polarizabilities $\alpha_p(i\omega)$ defined as

$$\alpha_p(i\omega) = \frac{\alpha_p^0}{1 - \left(i\omega/\omega_p\right)^2} \tag{2.39}$$

where ω_p is the characteristic excitation frequency for atom p. For the set of QHOs it is possible to write a Hamiltonian, where every QHO fluctuates within an isotropic harmonic potential

$$U(\vec{x}_p) = \frac{1}{2} m_p \omega_p^2 \vec{x}_p^2 \tag{2.40}$$

with dipole moments $\vec{d}_p = q_p \vec{x}_p$, where q_a is the charge of each oscillator. The Hamiltonian also incorporates dipole-dipole interaction terms and has the form

$$\hat{H}_{MBD} = -\frac{1}{2} \sum_{p=1}^{N} \frac{\nabla_{x_p}^2}{m_p} + \frac{1}{2} \sum_{p=1}^{N} m_p \omega_p^2 \vec{x}_p^2 + \sum_{p>q}^{N} \vec{d}_p^{\dagger} \mathbf{T_{pq}} \vec{d}_q \tag{2.41}$$

where $\mathbf{T_{pq}}$ is the dipole-dipole interaction tensor. The first term of the Hamiltonian represents the kinetic energy, the second term represents the potential energy of individual QHOs, while the third term represents dipole-dipole couplings. To compute the long-range correlation energy, the total energy of isolate oscillators needs to be subtracted from the total energy of the coupled ones (to cancel the contribution from kinetic and potential energy of QHOs). This leads to the following final equation

$$E_{MBD} = \frac{1}{2} \sum_{p=1}^{3N} \sqrt{\lambda_p} - \frac{3}{2} \sum_{p=1}^{N} \omega_p \tag{2.42}$$

where λ_p are the eigenvalues of $\hat{H}_{MBD}$.

A further model deserving attention and that has been employed in the present work is an adaptation of the TS model for solid surface systems. The model, named

TS^{surf}, was developed by Ruiz et al. [82, 83] to describe hybrid inorganic-organic systems. It is based on the TS [74] method combined with the Lifshitz-Zaremba-Kohn theory for the vdW interactions between an atom and a solid surface. Incorporation of the Coulomb screening effect within a bulk is achieved by rescaling the C_6 and R_{vdW} parameters of bulk atoms. Inclusion of screening effects can lead to a reduction of C_6 parameters up to a factor of 4 [82].

As it follows from the macroscopic theory developed by Lifshitz [84], the interaction between atom a and solid B is described by the following potential

$$E_{vdW} \simeq -\frac{C_3^{aB}}{(Z - Z_0)^3} \tag{2.43}$$

where Z is the distance between the atom and the surface while C_3^{aB} is the vdW coefficient. Later, Zaremba and Kohn [85] derived a formula,that connects C_3^{aB} with the polarizability of atom a ($\alpha_a(i\omega)$) and dielectric function of the solid B ($\epsilon_B(i\omega)$)

$$C_{3,LZK}^{aB} = \frac{\hbar}{4\pi}\int_0^\infty d\omega \alpha_a(i\omega)\frac{\varepsilon_B(i\omega) - 1}{\varepsilon_B(i\omega) + 1} \tag{2.44}$$

$C_{3,LZK}^{aB}$ can be related to $C_{6,LZK}^{ab}$ coefficients as follows [86]

$$C_{3,LZK}^{aB} = n_s\left(\frac{\pi}{6}\right)C_{6,LZK}^{ab} \tag{2.45}$$

where n_s is the number of atoms per unit volume in the bulk. Having computed the $C_{6,LZK}^{ab}$ it is now possible to obtain the α_{LZK}^b and $C_{6,LKZ}^{bb}$ parameters solving the system of two equations (like Equation 2.34) for any two atoms a (H, C, Ne, etc.) and solid B[14]. The R_{LZK}^b is then computed by the ratio cubic root of static dipole polarizabilites:

$$R_{LZK}^b = \left(\frac{\alpha_0^b}{\alpha_{0,free}^b}\right)^{1/3} R_{0,free}^b \tag{2.46}$$

where $R_{0,free}^b$ is the vdW radius of the free atom b. From this point the screened

[14]Suppose those screened parameters need to be computed for a Cu atom. We chose C and Xe atoms as references. Using the procedure described before, $C_{6,LZK}^{C-Cu}$ and $C_{6,LZK}^{Xe-Cu}$ can be obtained and two equations 2.34 can be written for them. Since the parameters of "free" C and Xe atoms are already known, the remaining unknown quantities are: α_{LZK}^b and $C_{6,LKZ}^{bb}$, which can be found trivially.

parameters[15] α^b_{LZK}, $C^{bb}_{6,LKZ}$ and $R^b_{0,free}$ are used in the standard TS or MBD schemes in place of the initial "free atom" parameters, establishing the so-called TS^{surf} or MBD^{surf} methods .

It has been demonstrated that the TS^{surf} method provides very accurate adsorption energy and adsorption heights of PTCDA[16], benzene and Xe atom on different substrates including Au(111), Pt(111), Ag(111) [82, 83]. It was also successfully used in the current work to unravel the origin of the striking enantioselectivity observed for the PdGa intermetallic compound (for more details see chapter 3).

2.4 DFT in HPC codes

Several approaches have been proposed in the literature to efficiently implement DFT in codes suitable for high-performance computing (HPC). To discuss those approaches, it is helpful to recall the set of equations that one aims to solve within the KS-DFT framework

$$\left[-\frac{1}{2}\nabla^2 + [\phi(\vec{r}) + \mu_{xc}(n)]\right]\psi_i(\vec{r}) = \epsilon_i\psi_i(\vec{r}) \tag{2.47}$$

So far, the actual representation of the KS orbitals $\psi_i(\vec{r})$ was not considered in this work. In practice they are defined as a linear combination of basis functions $\phi_j(\vec{r})$

$$\psi_i(\vec{r}) = \sum_\alpha c_{i\alpha}\phi_\alpha(\vec{r}) \tag{2.48}$$

The usage of a basis set allows to map the solution of a continuous problem into a task of finding the $c_{i\alpha}$ parameters. The choice of the basis set to a very large extent defines the numerical algorithms that will be used to solve the KS equations. Therefore, this choice stays at the very core of the development of a particular DFT code. One could distinguish two main types of basis sets: **plane-wave** and **atomic** orbitals.

Plane-wave basis sets define $\phi_\alpha(\vec{r})$ as $\frac{1}{\sqrt{\Omega}}e^{i\vec{G}_\alpha\vec{r}}$, where $\vec{G}_\alpha$ is the vector of reciprocal lattice. Plane-waves are a good choice from many different points of view. First of all, they are intrinsically orthogonal, which can used to accelerate the numerical

[15] The screened parameters inherit the effects from the many-body collective response of the substrate electrons.

[16] Perylenetetracarboxylic dianhydride.

algorithms. Plane-waves do not depend on the atomic positions. To improve the plane-waves basis, one needs to only increase their quantity, which is straightforward. They are easy to use and do not require optimization of a basis set for each particular atom. Plane-waves, however, intrinsically assume periodicity of the system to break which additional methods, such as Martyna-Tuckerman [87] should be employed. Plane-waves do not allow to separately improve the description of one particular region of the system. All (even empty) regions have to be populated with the same density of plane-waves, which lowers the computational efficiency of the approach. One cannot use them to represent the wave functions located closer to the nuclei, since those wave functions vary too rapidly. This problem is typically solved by introducing pseudopotentials.

In general form, **atomic orbitals** can be expressed as follows: $\phi(\vec{r}) = R(|\vec{r}_I|)Y(\hat{r}_I)$, where $R(|\vec{r}_I|)$ is the radial part of the function and $Y_{lm}(\hat{r}_I)$ is a spherical harmonic of the order l with angular momentum m. Depending on the function used for the radial part, atomic orbitals can be classified as Gaussians orbitals ($e^{-\alpha r^2}$), Slater-type orbitals ($e^{-\zeta r}$) or numerical orbitals. Atomic orbitals are typically very efficient, since they require only a small number of functions as they are (typically) centered at the position of atomic cores. The numerical complexity to describe orbitals close to the core is not as large as in the case of plane-waves. Also, they do not populate the vacuum as plane-waves do. Contrary to plane-waves, atomic orbitals do not imply system periodicity. However, they are not orthogonal and they suffer from basis set superposition error. And there is no unique way to enlarge the basis set, a simple approach of just adding many functions does not work in that case.

Plane-wave basis sets are implemented in CPMD [88], Quantum Espresso [89] and ABINIT [90] codes; SIESTA [91] and FHI-AIMS [92, 93, 94, 95] codes adopted atomic basis sets; a mixed approach that combines atomic basis sets and plane-waves is used in CP2K [96, 97].

The numerical approach employed in CP2K to solve DFT problems, known as "Quickstep", is described in a work by VandeVondele and coworkers [98]. The mixed Gaussian/plane-wave basis allows on one hand to benefit from the algebra of sparse matrices and, on the other hand, to efficiently perform operations in Fourier

space. The method can achieve a performance corresponding to an almost linear scaling (with respect to system size) in the case of large atomistic models [59]. In Quickstep, a Gaussian basis is used to represent the KS-orbitals to compute the electronic kinetic energy and the exchange-correlation energy in real space, while the plane-waves basis is used to represent the charge-density of the system to solve the Coulomb integral. In the Quickstep approach, the total charge density of the system when represented through a Gaussian basis set takes the form:

$$n(\vec{r}) = \sum_{i}^{occ} |\psi_i(\vec{r})|^2 = \sum_{i}^{occ} \psi_i^*(\vec{r})\psi_i(\vec{r}) = \sum_{i}^{occ}\sum_{\alpha,\beta} c_{i\alpha}^* \phi_\alpha^*(\vec{r}) c_{i\beta}\phi_\beta(\vec{r}) = \sum_{\alpha,\beta} \phi_\alpha^*(\vec{r})\phi_\beta(\vec{r}) \sum_{i}^{occ} c_{i\alpha}^* c_{i\beta} = \sum_{\alpha,\beta} \phi_\alpha^*(\vec{r})\phi_\beta(\vec{r}) P_{\alpha\beta} \tag{2.49}$$

where $P_{\alpha\beta}$ is an element of a so-called density matrix. The Gaussian-based density $n(\vec{r})$ is further represented on a 3-dimensional real space grid. Since the mapping on a grid involves several approximations, the new "on-grid" density is denoted as $\tilde{n}(\vec{r})$. $\tilde{n}(\vec{r})$ can be further represented in the reciprocal space using the fast Fourier transform (FFT)

$$\tilde{n}(\vec{r}) = \frac{1}{\Omega}\sum_{\vec{G}} \tilde{n}(\vec{G}) \exp(i\vec{G}\vec{r}) \tag{2.50}$$

where Ω is the volume of the simulation cell and $\vec{G}$ are the reciprocal lattice vectors. Such representation allows to efficiently compute the electrostatic potential solving Poisson's equation[17]

[17] The Hartree potential, is solved through the Fourier counterparts of $\tilde{V}(\vec{r})$ and $\tilde{n}(\vec{r})$

$$\tilde{V}(\vec{r}) = \frac{1}{\Omega}\sum_{\vec{G}} \tilde{V}(\vec{G}) e^{i\vec{G}\vec{r}} \tag{2.51}$$

$$\tilde{n}(\vec{r}) = \frac{1}{\Omega}\sum_{\vec{G}} \tilde{n}(\vec{G}) e^{i\vec{G}\vec{r}} \tag{2.52}$$

that allow writing

$$\nabla^2 \tilde{V}(\vec{r}) = \frac{1}{\Omega}\sum_{\vec{G}} -\vec{G}^2 \tilde{V}(\vec{G}) e^{i\vec{G}\vec{r}}$$

Substitution of $\nabla^2\tilde{V}(\vec{r})$ and $\tilde{n}(\vec{r})$ in the Poisson's equation with its representation in the Fourier spaces gives

$$\nabla^2 \tilde{V}(\vec{r}) = \frac{1}{\Omega}\sum_{\vec{G}} -\vec{G}^2 \tilde{V}(\vec{G}) e^{i\vec{G}\vec{r}} = -\frac{1}{\Omega}\sum_{\vec{G}} \frac{\tilde{n}(\vec{G})}{\varepsilon_0} e^{i\vec{G}\vec{r}} = -\frac{\tilde{n}(\vec{r})}{\varepsilon_0}$$

$$\nabla^2 V(\vec{r}) = -\frac{n(\vec{r})}{\varepsilon_0} \tag{2.53}$$

Operations in Fourier space, including the Fourier and back-Fourier transform (performed by means of efficient FFT libraries) and the sum in real space (on the grid points i), scale linearly with the system size and are thus a "solved problem" on parallel computers. Once the electrostatic potential is computed, the electrostatic energy can be found as the sum of $\tilde{V}(\vec{r}_i)\tilde{n}(\vec{r}_i)$ at every grid point

$$E_{Coulomb} = \sum^{grid} V(\vec{r}_i)n(\vec{r}_i) \tag{2.54}$$

Despite the fact that the CP2K code allows a so-called "all electrons" approach to compute the properties of a system, given the large size of the atomistic models studied in this work, "pseudopotentials" [99] were used to represent the non-valence orbitals.

The core electrons of each atom that are not influenced by the chemical bonding are considered frozen and are not included explicitly in the calculation. The valence electrons are treated explicitly while the nuclei and the frozen electrons are treated via an artificial potential that meets the following requirements for single atom properties: (i) the energy levels obtained for the all electron atom match the energy levels of the valence electrons of the pseudoatom; (ii) the wave functions of the electrons of the pseudoatom, outside a region close to the nucleus and defined by a parameter cutoff radius, match the wave fucntions of the corresponding electrons in the all electron atom.

Both properties grant the so-called transferability of the pseudopotential, that is the capability to correctly reproduce materials properties for different chemical

which allows to obtain very useful relationship

$$\tilde{V}(\vec{G}) = \frac{\tilde{n}(\vec{G})}{\vec{G}^2\varepsilon_0}$$

and finally provides a simple connection between electrostatic potential and the representation of charge density in the Fourier space

$$\tilde{V}(\vec{r}) = \frac{1}{\Omega}\sum_{\vec{G}} \frac{\tilde{n}(\vec{G})}{\vec{G}^2\varepsilon_0} e^{i\vec{G}\vec{r}}.$$

environments of the pseudoatom. The introduction of a cutoff radius (property (ii)) allows to remove the sharp oscillations that the wave function has in the region close to the nucleus due to the orthogonality condition to the inner electrons of the system. A smooth function can be represented with a basis set that is not exceedingly large.

2.5 Materials properties from DFT

So far the discussions in this chapter were only focused on how to access the electron density and total energy of a system. Even though this information by itself is interesting to obtain, it is important to derive from it the measured physical, thermodynamical and spectroscopic quantities. The starting point of such an investigation is the definition of an atomistic model of a system that, with a limited number of particles[18], is capable of catching the physical and chemical aspects of the object under study. Recent implementations of DFT, as the one described in the previous section for the CP2K code [96, 97], allow, by means of standard HPC infrastructures, to deal comfortably with systems including a few thousands of atoms.

One of the basic questions computational materials science can answer is "what is the equilibrium geometry" of the system represented by a model. Having access to the ground-state energy of the model, and considering valid the hypothesis of the Born-Oppenheimer (BO) approximation [51][19], it is straightforward to determine the forces acting on the atomic nuclei due to the system of electrons. Considering the nuclei positions $\vec{R}_i$ as parameters to the functional $E[n(\vec{r})]$, the Hellman-Feynman theorem [49, 50] allows to compute the forces $\vec{F}_i$ acting on $\vec{R}_i$ as

$$\vec{F}_i = -\langle \Psi | \frac{d\hat{H}}{d\vec{R}_i} | \Psi \rangle \tag{2.55}$$

Applying efficient minimization schemes such as BFGS [100, 101, 102, 103] to the 3N degrees of freedom represented by the N atoms of the model, the local minima of the energy of the system (as a function of the atomic coordinates) can be identified.

[18] The limit is dictated by the computational cost at a specific "level of theory".

[19] Born-Oppenheimer (or adiabatic) approximation assumes absence of the crossing of the potential energy curves, corresponding to the different electronic states. Those crossings create non-adiabatic effects which are neglected by the BO approximation.

Defining the basis of *ab inito* thermodynamics [104], the optimized geometries (and their corresponding energies) are, typically, considered as a starting point for further investigations. Moreover, by itself they can already be employed to access some relevant thermodynamic quantities. For example, the adsorption heat (which can be measured experimentally) for a molecule 'a' adsorbed on the substrate 'B' can be computed[20] as follows

$$q_{ads} = \Delta H \approx \Delta U \approx E_{aB} - E_a - E_B \tag{2.56}$$

where ΔH is the change of enthalpy, ΔU is the change of total energy, E_{aB} is the total *static* energy of the molecule/substrate complex and E_a/E_B are the total *static* energies of the molecule and substrate separately.

In the transformation processes, where entropy plays an important role, it is not enough to consider the total energy change. In that case, the change of the *Gibbs energy* provides the direction of the process. Such problems can be ideally accessed performing molecular dynamics simulations governed by the *ab initio* forces or following minimum energy path approaches as will be discussed in section 2.7. Infrared spectrum, zero-point energy (ZPE) and vibrational entropy can be derived from the phonon spectrum, to compute which the finite differences approaches or the formalism of linear response can be employed [105]. In some cases for example, when comparing the energies of surfaces with different stoichiometry [12] the atomic chemical potentials are employed to describe the interaction of an atom with the outer environment. The investigation of surface stability goes beyond the scope of the current study therefore the field of chemical potentials will not be addressed here.

The way to access the electronic properties of a system is less direct. Looking back at the eigenvalues ϵ_i of Equation 2.47, it can be demonstarted that the eigenvalue of the highest occupied KS orbital corresponds to the Fermi energy of the system. However, no physical meaning can be rigorously attributed to the remaining

[20] The adsorption energy is provided by the change of enthalpy if the experiment is done under constant pressure conditions. The neglecting $\Delta(PV)$ term allows to write $\Delta H \approx \Delta U$ (note, if experiment is done at constant volume it is correct to directly write $q_{ads} = \Delta U$). ΔU includes both: static and vibrational contributions: ΔE_{stat} and ΔE_{vib}. Assuming that vibrational contributions do not change significantly during the adsorption process it is possible to write: $\Delta U = \Delta E_{stat}$.

eigenvalues[21]. Incidentally, for a large class of systems the KS eigenvalues provide a reasonable approximation for the single particle excitation energies giving access, for example, to the electronic bandstructure of a bulk system. For the discussion about interpretation of the energy of KS orbital energies the reader is referred to the paper by Chong et al. [107].

It has to be remarked that, independently from the approximation adopted in the expression of the exchange correlation potential, even a fundamental quantity such as the transport gap (or the HOMO-LUMO gap in a molecular system) can be considerably underestimated by DFT. Moreover, the most common expressions for the exchange correlation potential fail to compensate for the self interaction energy introduced by the Coulomb term in Equation 2.6, causing a systematic error whose relevance is non negligible for systems characterized by charge localization.

Still, the KS eigenvalues and the shape of the KS orbitals usually provide a basis for a reasonable description of spectroscopic properties such as scanning tunneling microscopy (STM) or scanning tunneling spectroscopy (STS) for molecules adsorbed on a substrate. Moreover, KS eigenvalues and orbitals are usually employed as a starting point for higher levels of theory such as the GW approach [108].

2.6 Computation of effective point charges

"Atomic charges" is not a well defined concept, taking into account the distributed nature of electron density. However, it is a very useful tool for understanding the chemical/physical properties of molecules. An approach of assigning point charges to atoms was originally used by Momany [109], who fitted them to reproduce the electrostatic potential around the molecule by minimizing the root-mean-square deviation between DFT-computed electrostatic potential $V_{DFT}(\vec{r}_i)$ on a certain grid/set of N points appropriately chosen in space and electrostatic potential formed by the point charges q_j

$$RMSD = \frac{1}{N}\sum_{i}^{N}\left(V_{DFT}(\vec{r}_i) - k\sum_{j}\frac{q_j}{|\vec{r}_{ij}|}\right)^2 \tag{2.57}$$

[21] Please note that in the context of the time-dependent DFT a physical meaning to the unoccupied eigenvalues **can** be attributed. As it was shown by Gritsenko and Baerends [106] the energy difference between virtual and occupied KS states approximate the electron excitation energy.

where N is the number of points in 3D space used for fitting, $\vec{r}_i$ is the vector pointing to the point i, $\vec{r}_{ij}$ is the vector pointing from the charge j to the point i, k is Coulomb's constant.

The charges computed by this method reproduce intermolecular interactions very well. However, they have several disadvantages: conformational dependence, exaggerated polarization of molecular bonds resulting in wrong intramolecular interactions. To overcome the problem of transferability of charges and conformational dependence, Bayly et al. introduced a restraint on the atomic charges [110]. A penalty function was used in order to restrain the atomic charges to more attenuated values. This modified approach to fit the atomic charges of a system is called restrained electrostatic potential (RESP).

2.7 Characterization of molecular transformations at surface

The computer simulation of complex systems at the atomistic level is becoming increasingly useful in solid state physics, chemistry and biophysics but still shows shortcomings with respect to more empirical methods due to the fact that an atomistic description is expensive. The class of methods that aim to simulate trajectories of a system integrating the equations of motion go under the name of molecular dynamics (MD) [111]. One of the bottlenecks for extracting the system properties using such methods resides in the fact that the dynamics of realistic systems encompasses a large variety of characteristic times, and the integration of the equations of motion requires a time step that is shorter with respect to the period of the fastest involved dynamics of the system. This leads to an upper limit for the time scales that can afforded by standard molecular dynamics simulation: of the order of one nanosecond for *ab initio* simulations and of one microsecond for classical approaches. Unfortunately, most of the interesting phenomena occurring in complex systems happen on a timescale that is orders of magnitude larger: protein modifications, chemical reactions, structural phase transitions can be thus classified as *rare events*, either because, albeit being intrinsically fast, they have a very low probability (reactions with high energy barrier) or because the underlying potential energy surface implies a slow and diffusive behavior of the dynamical trajectories.

In order to overcome the difficulties related to the too large timescales involved in the sampling of a rare event, a coarse-grained model instead of an all-atom description can be adopted in a simulation. The choice of the variables to coarse requires however a profound knowledge of the systems that is not a priori available. Keeping the atomistic description is possible by exploiting methods aimed at the acceleration of rare events. Following Laio and Gervasio [112], one can classify the methods into four categories: the ones aimed at enhancing the sampling as a function of some preassigned collective variables; the ones aimed at exploring the transition mechanisms of a reaction and finding reactive trajectories; the ones for localizing saddle points starting from potential energy surface minima, and the ones where the phase space is explored at different temperatures and statistically based exchanges are performed during the simulation.

Metadynamics encompasses several features of these four classes of methods into a unified strategy for accelerating rare events. The algorithm is based on the dimensional reduction of the problem into a set of collective variables. The dynamics in the space of the chosen variables is modified by adding a history-dependent term in the potential made up by a sum of Gaussians that discourage the system from revisiting configurations with the same values of the collective variables. The original paper by Laio and Parrinello was entitled "Escaping free energy minima" [113], a title that well summarizes one of the two main uses of this method, namely to evolve from a stable basin to another, unknown, one via the lowest possible saddle point: starting from a minimum of the potential, a walker (a system subject to molecular dynamics evolution) increases its energy level in the potential energy well through the successive deposition of sand (Gaussians), until the sum of the already deposited sand (the sum of all Gaussians) allows the walker to pass the first available transition state and the minimum is escaped. The other use of the metadynamics method is to provide an estimator for the free energy of the system.

The key to the method is an appropriate choice of the collective variables, which should demonstrate the following properties: (i) they should distinguish between initial, intermediate and final states; (ii) they should describe the slow events underlying the process under study; (iii) their number should not be too large, to allow to

fill the free energy basins in a reasonable time. If a relevant collective variable is neglected, it can be shown with examples that the barrier height can be overestimated and/or the reconstructed potential energy surface can show hysteretic behavior.

Examples of collective variables are an angle, a distance, a coordination number, the potential energy, the gyration radius.

To summarize, $U(\vec{r})$ is the potential energy of a system (for example a molecule) with $\vec{r}$ the coordinates of the atoms. It is convenient to define so-called "collective variables" $C(\vec{r})$, functions of the atomic coordinates, that highlight the relevant atomistic details during a transformation of the system. They are selected to investigate a particular transformation process and it reflects the advancement of the system with respect to this process.

Therefore, applying a penalty function that increases the potential $U(\vec{r})$, as a function of a collective variable, allows to push the system to explore the potential along this collective variable. This is achieved through so-called "history dependent" potential which gets a penalty function every N_G steps with $\tau_G = \tau N_G$, where τ is a step of MD

$$U(C(\vec{r}),\vec{r},t) = U(\vec{r}) + \omega \sum_{t'<t, t'=\tau_G, 2\tau_G, \ldots} \exp\left(-\frac{(C(\vec{r}) - c(t'))^2}{2s^2}\right) \tag{2.58}$$

where ω is the Gaussian height, s is its width $c(t')$ is the value of collective variable at time t'. It can be proven that the following limit converges to the free energy profile $F(C)$ along the collective variable $C(\vec{r})$:

$$F(C) = -\lim_{t\to\infty} [U(C(\vec{r}),\vec{r},t) - U(\vec{r})] \tag{2.59}$$

Nevertheless, to escape a free energy minimum by means of a metadynamics simulation, it requires at least a number of simulation steps proportional to the ratio between the volume of a free energy basin and the volume of the biasing Gaussians. For an event requiring a barrier of 1.5 eV, supposing the basin has an extension of 2 in units of the collective variable, we can estimate a number of steps in the order of 150000. This makes the application of metadynamics unfeasible where the forces of ions have to be derived by *ab initio* methods applied to large scale systems such as

the ones used to represent molecules on surfaces.

In this work, metadynamics is applied to study the transformation of a molecule called 9-ethynylphenanthrene (in particular, to explore qualitatively the interconversion mechanism) adsorbed on a PdGa surface. In this latter case, one can simplify the problem computing activation barriers from a potential energy profile (that would resemble a free energy profile at zero temperature). Such a method is described in the next section.

2.8 Nudged elastic band method

To introduce the nudged elastic band (NEB) method, it is useful to recall the basic concepts of transition state theory (TST)[22]. TST assumes the existence of a hypersurface of dimensionality D in the phase space with two divided regions corresponding to reactants and products [115]. The dividing hypersurface of dimensionality $D-1$ is defined in such a way that the reacting trajectory going from the initial to final state crosses it only once. The reaction is assumed to be electronically adiabatic, so the Born-Oppenheimer approximation [51] remains valid.

If the reaction happens at a crystal surface at temperatures significantly lower than melting temperature, the harmonic approximation can be applied to TST (abbreviated as hTST). The transformation process within hTST can be viewed as the crossing of one of the lowest saddle points surrounding the potential energy basin of the initial state. The reaction rate k is then provided by Arrhenius' equation

$$k^{hTST} = \Gamma^{hTST} e^{\left(-\frac{\Delta E^{\#}}{k_B T}\right)} \tag{2.60}$$

where Γ^{hTST} is the *attempt frequency* [23], $\Delta E^{\#}$ is the reaction energy barrier (see Figure 2.2). Within hTST, Γ is computed as

$$\Gamma^{hTST} = \frac{\prod_i^D \nu_i^{init}}{\prod_i^{D-1} \nu_i^{\#}} \tag{2.61}$$

[22] For more comprehensive overview of the current state of the TST theory the reader is refereed to the following article [114].

[23] Can be viewed as the number of movements per second along the reaction coordinate.

where ν_i^{init} are the normal mode frequencies at the initial state, $\nu_i^{\#}$ are the normal mode frequencies at the saddle point excluding the imaginary frequency defining the saddle point. In this work the Γ^{hTST} was not computed directly, since it would be prohibitive, but the value of 10^{13}, standard for processes at surfaces [116], was used instead to estimate the reaction constant. To obtain the $\Delta E^{\#}$ one needs first to find the *minimum energy path* (MEP) connecting the initial and final states of the reaction. Since the MEP can have several saddle points, only the highest one will determine the reaction rate (this is true only if Boltzmann distribution is reached for the metastable points on the MEP). Therefore, it is necessary to obtain a good estimation of the MEP in order to find the highest saddle point that will be used in the computation of the reaction rate.

NEB serves exactly this purpose. It provides an efficient way of finding the MEP between the initial state (**A**) and the final state (**B**) [117]. Initially NEB was applied by Mills and Jónsson to study the dissociative adsorption of H_2 on a Cu(110) surface [118]. The idea behind the method can be described as follows. The initial approximation to the MEP is provided by a set of replicas of the system [$\mathbf{A}=\mathbf{R_0}$, $\mathbf{R_1}$, ..., $\mathbf{R_N}=\mathbf{B}$] ($\mathbf{R_i}$ denotes all coordinates of an image **i**) that are connected by springs forming a smooth path from **A** to **B**. A feature that is specific to the NEB method is a splitting of the "true force" ($\vec{F}_i$) and spring force ($\vec{F}_i^s$) into two components: perpendicular ($\perp$) and parallel ($\parallel$) to the tangent of the reaction path. The minimization procedure acts in correspondence to the force F_i obtained as

$$\vec{F}_i = {}^{\perp}\vec{F}_i + {}^{\parallel}\vec{F}_i^s \tag{2.62}$$

where the perpendicular component of true force ${}^{\perp}\vec{F}_i$ is computed

$${}^{\perp}\vec{F}_i = -\nabla {}^{\perp}E(\mathbf{R_i}) = -\nabla E(\mathbf{R_i}) + (\nabla E(\mathbf{R_i}) \cdot \hat{\vec{\tau}}_i)\hat{\vec{\tau}}_i \tag{2.63}$$

and the parallel component of spring force is computed as

$${}^{\parallel}\vec{F}_i^s = k\left[(|\mathbf{R_{i+1}} - \mathbf{R_i}| - |\mathbf{R_i} - \mathbf{R_{i-1}}|) \cdot \hat{\vec{\tau}}\right]\hat{\vec{\tau}} \tag{2.64}$$

where $\hat{\vec{\tau}}_i$ is the normalized local tangent at the replica i. Such forces can be used

to reduce and minimize the total energy of the path. Figure 2.2 presents a pictorial view of the concept of an NEB calculation. The figure describes a simple mechanical model of a ball passing over a hill of a certain height where its potential energy at the top is increased by unknown $\Delta E^{\#}$ with respect to the initial position **A**.

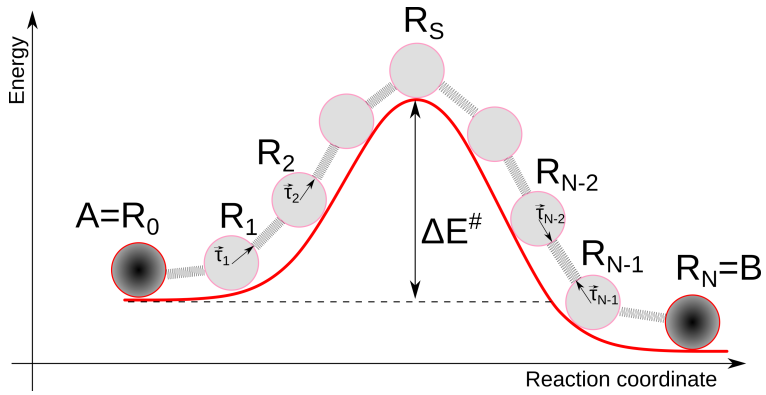


FIGURE 2.2: A pictorial representation of the NEB approach. A simple mechanical model, where replicas $\mathbf{R_i}$ are represented by balls connected with identical springs. $\vec{\tau}_i$ is the local tangent at the replica i.

In the case of molecular systems, due to the internal degrees of freedom the picture becomes more complex, however the idea remains pretty much the same.

Although the original NEB approach was widely used, it suffers from a noticeable drawback: the replica $\mathbf{R_i}$ tends to avoid the saddle point on MEP, therefore the $\Delta E^{\#}$ can only be obtained via interpolation. To overcome this problem, Henkelman et at. proposed a so-called "climbing image" NEB method [119]. In this approach, a rigorous convergence to a saddle point ($\mathbf{R_S}$ in Figure 2.2) can be achieved. After several iterations, climbing image NEB identifies the replica with the highest energy $\mathbf{R_{max}}$. From that moment, the force $\vec{F}_{max}$ acting on $\mathbf{R_{max}}$ becomes

$$\vec{F}_{max} = -\nabla E(\mathbf{R_{max}}) + 2(\nabla E(\mathbf{R_{max}}) \cdot \hat{\vec{\tau}}_{max})\hat{\vec{\tau}}_{max} \tag{2.65}$$

while the forces on the other images $\mathbf{R_i}$ remain as in Equation 2.62. Such a modification enforces the image $\mathbf{R_{max}}$ to move "...up the potential energy surface along the elastic band and down the potential surface perpendicular to the band" [119]. Once the optimization is converged (implying that $\mathbf{R_{max}} = \mathbf{R_S}$), the energy barrier can be computed as

$$\Delta E^{\#} = E(\mathbf{R_{max}}) - E(\mathbf{R_A}) \tag{2.66}$$

For more rigorous mathematical derivation and generalization of the NEB (with the so-called "string" method) the reader is referred to the articles by van den Eijden E et al. [120, 121].

2.9 Scanning tunneling microscopy simulations

Scanning tunneling microscopy (STM) is a method which allows to obtain atomic resolution images of metallic or semiconductor materials and is based on the quantum tunneling effect. The basic setup is shown in Figure 2.3. An induced electrostatic potential difference (or bias voltage, V_{bias}) between tip and sample enforces electrons to tunnel, causing the tunneling current which is then plotted as a function of the tip's position.[24]

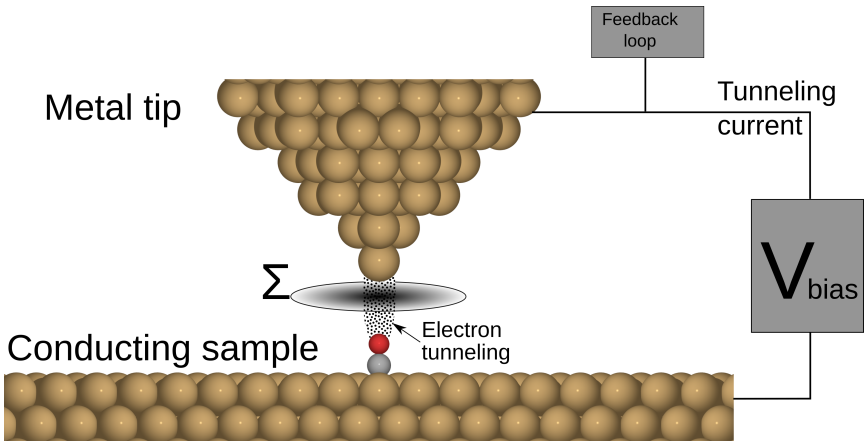


FIGURE 2.3: A schematic representation of the STM setup: the sharp metallic tip and conducting sample are connected with a wire, applying bias voltage (V_{bias}) between them. The tunneling current, which arises due to the electron tunneling between the tip and sample, is further detected and plotted as a function of the tip's position.

Despite the fact that the STM technique has been improving continuously since its discovery and is capable now of resolving sharply atomistic details, comparison between measured and simulated STM images is still the most efficient approach to understanding the chemical and electronic features of a system under investigation. The theoretical basis to simulate STM images dates back to the works of Bardeen [122], initially developed to describe the tunneling current between two

[24] This is one possible approach, called "constant height mode", where the tip-sample distance is constant while the tunneling current changes as a function of the tip's position. Another possible way of doing the STM measurements is adapting the tip's height in order to keep the current fixed. This approach is called "constant current mode".

electrodes. Bardeen's theory is still widely adopted in computational tools to simulate STM. According to Bardeen, the tunneling current between the tip and conducting sample can be written as

$$I(V) = \frac{2\pi e}{\hbar} \sum_{\mu,\nu} \delta(E_\mu - E_\nu + eV) |M_{\mu\nu}|^2 \rho(E_\mu)\rho(E_\nu) \left[f(E_\mu) - f(E_\nu) \right] \tag{2.67}$$

where μ and ν denote the electronic states of the tip and of sample respectively in a single particle picture, $f(E)$ is the Fermi-Dirac distribution function[25], $\rho(E_{\mu/\nu})$ is the density of states of the tip and sample, V is the applied bias voltage, $E_{\mu|\nu}$ are the energies of states $\psi_{\mu|\nu}$ and $M_{\mu\nu}$ is the tunneling matrix describing the electron tunneling between states ψ_μ and ψ_ν. $M_{\mu\nu}$ is defined in the following way

$$M_{\mu\nu} = \frac{\hbar^2}{2m} \int_\Sigma \left(\psi_\mu^* \nabla \psi_\nu - \psi_\nu \nabla \psi_\mu^* \right) d\vec{S} \tag{2.68}$$

where the integral is extended to any separating surface Σ between the tip and sample (see Figure 2.3).

The approach outlined by Bardeen, where the electronic correlation does not dramatically influence the tunneling current, can be in many cases simplified following the well-known Tersoff-Hamann approximation (TH) [123]. The TH approximation starts from the assumption of a low temperature, so that the Fermi function in Equation 2.67 can be cosidered constant up to the Fermi level and zero beyond Fermi. Rewriting the term $\left[f(E_\mu) - f(E_\nu) \right]$ in Equation 2.67 as

$$\begin{aligned} f(E_\mu) - f(E_\nu) &= f(E_\mu) - f(E_\mu)f(E_\nu) + f(E_\mu)f(E_\nu) - f(E_\nu) \\ &= f(E_\mu)(1 - f(E_\nu)) - f(E_\nu)(1 - f(E_\mu)) \end{aligned}$$

reveals that the low temperature assumption allows to limit the sum in Equation 2.67 to the tip states in the energy range $[E_F - eV : E_F]$ and sample states in the energy range $[E_F : E_F + eV]$. A further assumption in the TH approach is "low bias voltage", meaning that the $\psi_{\mu|\nu}$ states are supposed to be unperturbed upon setting

[25] The Fermi-Dirac distribution function has the following form: $f(E) = 1/\left[1 + \exp\left(\frac{E-E_F}{k_B T}\right)\right]$, where E_F is the Fermi energy and k_B is the Boltzmann constant.

of the bias voltage. The tunneling current can now be found as follows

$$I(V) = \frac{2\pi e}{\hbar} \sum_{\mu,\nu} \delta(E_\mu - E_\nu + eV)|M_{\mu\nu}|^2 \rho(E_\mu)\rho(E_\nu) \tag{2.69}$$

where the μ goes through all the states in the energy range $[E_F - eV : E_F]$ and ν goes through the states in the energy range $[E_F : E_F + eV]$. To get rid of the tunneling matrix, which depends on the physical shape of the tip (generally unknown), TH approximation assumes the region of interest (e.g. at the tip apex), the tip's wave funtions is basically spherical. Assuming a s-orbital-like wave function for the tip, the TH approximation finally provides a simple formula for the tunneling current

$$I(V) = \propto \sum_{\nu} |\psi_\nu(\vec{r}_0)|^2 \rho(E_\nu - eV) \tag{2.70}$$

The STM image in the TH approximation is thus proportional to the local density of states of the sample system at point $\vec{r}_0$, and provides an intuitive way to interpret STM images.

The TH approximation has been successfully applied in a large variety of cases. Even experiments where the tip of the microscope is functionalized e.g. by means of a CO molecule, TH can be employed provided that the p character orbital of the functionalizing group is considered. In this latter case, Chen has shown [124] that it is enough to consider spatial derivatives of the substrate's wave functions at the position of the tip ($\vec{r}_0$).

A severe technical problem encountered in simulating STM images comes from the fact that at the typical experimental tip-sample distances, the mathematical description provided by standard DFT approaches for the decay of the single particle orbitals in the vacuum region is inappropriate (the decay is much faster compared to the expected exponential decay). Tersoff [125] proposed a method to extrapolate the tails of the wavefunctions starting from a plane parallel to the surface located in a region where the single particle orbitals are still properly described. For more details on the extrapolation procedure and its applicability, we refer the reader to the original article of Tersoff [123] and to the work by Gaspari and coworkers [126].

A practical toolkit to simulate STM images starting form KS orbitals and extrapolating the orbitals in a vacuum region far from the substrate is available from L. Talirz [127].

2.10 Noncontact atomic force microscopy simulations

Another kind of measurement that belongs to the scanning probe microscopy branch and can be modeled by modern simulation techniques is *atomic force microscopy* (or AFM, see introduction in section 1.5). Despite the striking quality of noncontact AFM (nc-AFM) images[26] most of the peculiar nc-AFM features require to be compared to a simulation in order to assign them a specific origin at the atomistic level. Fruitful employment of modeling can be appreciated by looking at various results obtained in the field. For example, using simulations the origin of high resolution in nc-AFM imaging was attributed to Pauli repulsion [128], distortion of the benzene ring was assigned to the spatial extent of its π-orbitals and the tilting of the CO molecule [129, 130, 131]. Additional efforts were required to understand the origin of the bond-like features present in some cases of hydrogen bonding [132]. It was demonstrated that the straight lines in nc-AFM images interpreted as "bonds" could also appear between two non-bonded atoms [133]. Simulations revealed that the nc-AFM method is sensible to the ridge of the potential energy between atoms [131, 33] rather then to increased charge density; thus, it can show hydrogen bonds as "sticks" but can also show bonds where no bonding exists.

In the experiment, the recorded signal corresponds to frequency shifts in nc-AFM cantilever oscillations (the period of an oscillation is typically around 1 - 100 microseconds). In the simulation, on the other hand, it is not possible to access that timescale, thus images are obtained processing the force acting on the AFM tip. According to Sader [134], the vertical component of the total force acting from the sample on the cantilever (F_z^{sam}) can be directly transformed into a frequency shift as follows

$$\frac{\Delta\omega}{\omega_0} = -\frac{1}{\pi a k}\int_{-1}^{1} F_z^{sam}(z+au)\frac{u}{\sqrt{1-u^2}}d \tag{2.71}$$

[26] Obtained in most of the cases for flat molecular systems.

where ω_0 is the resonant frequency of the unperturbed cantilever, $\Delta\omega$ is the frequency shift due to the interaction with the sample, a is the oscillation amplitude and k is the stiffness of the cantilever. Forces are the quantities that are accessible in both *ab initio* and classical models; however, since only the contribution from the sample enters Equation 2.71, this point has to be discussed in more detail. In the classical models, the separation of the sample contribution is straightforward, but in some cases (as for example in DFT) it cannot be done easily. If we consider the tip and sample together as an isolated system, then every atom of the tip (and of the attached CO molecule) experiences a force $\vec{F}_\alpha^{tip}$ that originates from interactions with other atoms of the tip and from interactions of the tip with the sample

$$\vec{F}_\alpha^{tot} = \underbrace{\sum_{i\neq\alpha}^{tip} \vec{F}_i}_{\vec{F}_\alpha^{tip}} + \underbrace{\sum_{j}^{sam} \vec{F}_j}_{\vec{F}_\alpha^{sam}} \tag{2.72}$$

Assuming that the presence of the sample has a negligible influence on the internal forces of the tip ($\vec{F}_\alpha^{tip}$), it is possible to replace the latter forces with the ones acting among the atoms of an isolated tip $\vec{F}_\alpha^{isol.tip}$ being exactly in the same geometry. This assumption allows to directly compute

$$\vec{F}_\alpha^{sam} = \vec{F}_\alpha^{tot} - \vec{F}_\alpha^{isol.tip} \tag{2.73}$$

Thus, the total force from the sample acting on the tip is simply the vectorial sum of $\vec{F}_a^{sub}$: $\vec{F}^{sub} = \sum_\alpha^{tip} \vec{F}_a^{sam}$. It is also possible to compute $\vec{F}_z^{sam}$ as a derivative of the total energy $\vec{F}_z^{sam} = -\frac{\partial E}{\partial z}$ with respect to the displacement of the tip along z direction. In principle, the latter approach is exact, but computing the numerical derivative introduces an error which depends on the step size, making this approach less interesting from a practical point of view.

Having discussed the issue of separating sample contribution to the force acting on the tip, it is now appropriate to discuss the existing models to compute forces emphasizing their strong and weak points. Speaking generally, since the vertical force significantly depends on the position of the functionalization molecule attached to the tip (for simplicity we consider a CO molecule in the following), an "ideal model"

should allow to relax its position according to the probe/sample interactions. Pauli repulsion [29, 128], dominant at close tip-sample distances and crucial for a high resolution, the vdW attraction and the electrostatic interaction have to be also included in the model [135]. From this perspective, vdW-inclusive DFT approximations seem to be the method of choice, and indeed it was used in several studies: to connect the "bond order" in chemical bonds and brightness of the "bond features" in nc-AFM images [30] , to explain the distortions of the benzene ring in the nc-AFM images [136, 131], to unravel the origin of the contrast which is interpreted as intermolecular bonds [131]. However, employing DFT to obtain nc-AFM images requires a prohibitive amount of computational resources, and therefore can only be done for small molecules (but still requires a significant amount of computational recourses).

Empirical approaches to the simulation of nc-AFM images are emerging as an alternative to the challenging DFT approach. The so-called ProbeParticle (PP) model proposed by Hapala et al. [33, 34] mimics the functionalizing molecule by a single particle Figure 2.4, which interacts with the sample through Pauli, van der Waals and Electrostatic interactions (as it was required for the "ideal model") properly parameterized. PP is attached to the end of the tip with a spring of stiffness k_{rad}, while the lateral movement of the PP is controlled by two lateral springs k_{lat} along X and Y directions. The metallic tip is represented as a fixed point which does not interact with the sample.

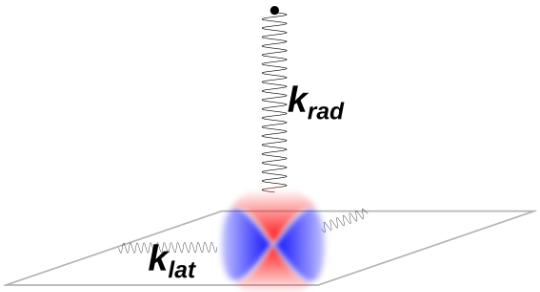


FIGURE 2.4: A schematic representation of the ProbeParticle model. The ProbeParticle is attached to the end of the tip with a spring of stiffness k_{rad}, while the lateral movements are controlled by k_{lat}. The charge density around the PP is represented by the red/blue cloud.

In the PP model, the Lennard-Jones potential is used to account for dispersion attraction and Pauli repulsion

$$V_{LJ}(r) = \varepsilon \left[\underbrace{\left(\frac{r_{min}}{r}\right)^{12}}_{\text{Pauli repulsion}} - \underbrace{2\left(\frac{r_{min}}{r}\right)^{6}}_{\text{vdW attraction}} \right] \tag{2.74}$$

where r is the distance between two atoms, r_{min} is the distance corresponding to the minimum on the energy profile, ε is the value of $V_{LJ}(r)$ at r_{min}. The electrostatic interaction is computed as a convolution of the Hartree potential (typically obtained from a DFT simulation) of the sample and an *ad hoc* charge density $n(\vec{r}', \vec{r})$ (where $\vec{r}'$ is the position on PP and $\vec{r}$ is a vector pointing to an arbitrary point in space) attributed to the PP

$$E_{El} = \int E_H(\vec{r}) n(\vec{r}', \vec{r}) d\vec{r} \tag{2.75}$$

where $E_H(\vec{r})$ is the Hartree potential originated by the sample and $n(\vec{r}', \vec{r})$ is the charge density with total charge Q attributed to the probe and $\vec{r}'$is the position of the PP

$$n(\vec{r}', \vec{r}) = \frac{Q}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{\vec{r}'-\vec{r}}{\sigma}\right)^2} \tag{2.76}$$

The computational cost of the PP approach is negligible, thus allowing to obtain promptly simulated nc-AFM images on a desktop computer. The drawbacks of the PP model are the non-physical origin of the parameters needed for the analytic form of the potential and the reliability itself of the force field.

Chapter 3

Adsorption of 9-EP molecules on PdGa:A(111) surface

Despite the fact that DFT [47, 48] has been established as a reliable standard in material science for the study of materials properties, accurate modeling of adsorption phenomena for organic molecules on metal surfaces on a first-principles level is still challenging, especially due to the large size of atomistic models needed to reliably describe the physics and chemistry of the system under investigation. Moreover, traditional density functional approximations (DFAs) fail to reproduce van der Waals (vdW) interactions that are crucial for the correct estimate of adsorption energies and geometries in adsorbate-substrate systems [138, 139]. In recent years, several vdW-inclusive DFAs have been developed, ranging from pairwise-additive models, over non-local functionals, up to many-body approaches [68, 69, 74, 140, 141, 76, 142]. In particular, it was shown that the inclusion of many-body dielectric screening effects leads to an improved description of geometries and energetics for molecule/surface systems [138, 139], supermolecular complexes [143], and molecular crystals [144, 145, 146].

In this chapter, following the introduction on the PdGa that was provided in section 1.3, we investigate the chiral properties of the (111) surface of this material. Specifically we refer to the recent work of Prinz et al. [9] on the adsorption of 9-ethynylphenanthrene on a PdGa(111)A:Pd_1 surface. By means of DFT-based

atomistic simulations, we unravel the origin of the experimentally observed striking enantioselectivity.

It turns out that the key ingredient to understanding the experimental evidence is the appropriate description of van der Waals interactions beyond the widely employed atomic pairwise approximations. A recently developed van der Waals-inclusive density-functional method, which encompasses dielectric screening effects, reveals the origin of the experimentally observed enantioselectivity and provides conclusive evidence of chiral recognition on a bimetallic surface driven by dispersion interactions. The incorporation of dielectric screening leads to a renormalization of the dispersion interaction range allowing for the appropriate weighting of the molecule-substrate interactions at intermediate distances between 2.5 and 5 Å. Such findings have implications for the structure and stability of complex organic/inorganic systems where dielectric screening effects are expected to be of general importance. We provide the first evidence of chiral recognition on a surface driven by dispersion interactions.

3.1 Adsorption geometry of 9-EP on a Pd_1 surface

To demonstrate the *R* and *S* adsorption geometries hypothesized in the experiments, we show in Figure 3.1 one of the experimental STM images measured by Prinz and coworkers. We focus here on the case of *R* and *S* enantiomers on a defect-free surface.

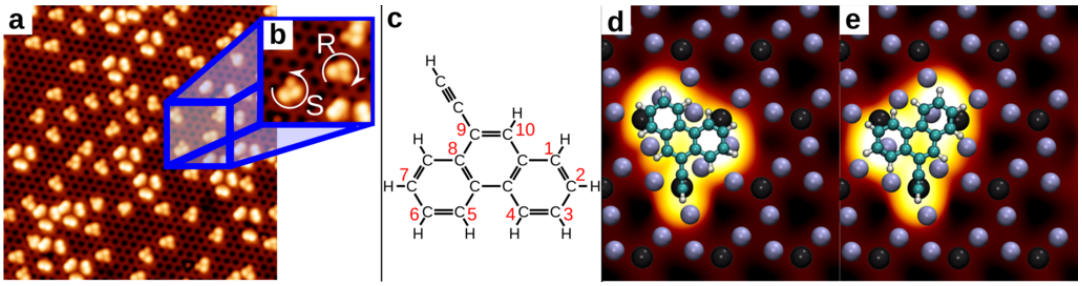


FIGURE 3.1: Experimental and theoretical STM images for the adsorption of 9-EP on PdGa:A(111)Pd_1. (a) *R* and *S* adsorbate on the PdGa:A(111)Pd_1 surface. (b) Zoom showing *R* and *S* surface enantiomers. (c) Chemical sketch of the 9-EP molecule laying in the *R* and *S* orientation. (d,e) Simulated STM images showing agreement with the experimental data. *Experimental images courtesy of Samuel Stolz and Roland Widmer from the nanotech@surface laboratory at Empa.*

In the experimental STM images, the 9-EP molecules present an asymmetric three-lobe morphology whose handedness can be inferred from the sense of rotation going from the largest to the smallest lobe (see Figure 3.1a,b). Prinz et al. [9] concluded that the 9-EP molecule is located on top of three isolated Pd atoms.

A detailed analysis of the STM patterns made it evident that two different adsorption sites can be distinguished for both enantiomers: R and R^* as well as S and S^*, where "starred" configurations are located on top of a Ga trimer from the second layer while the "non-starred" configurations are aside of it (Figure 3.2).

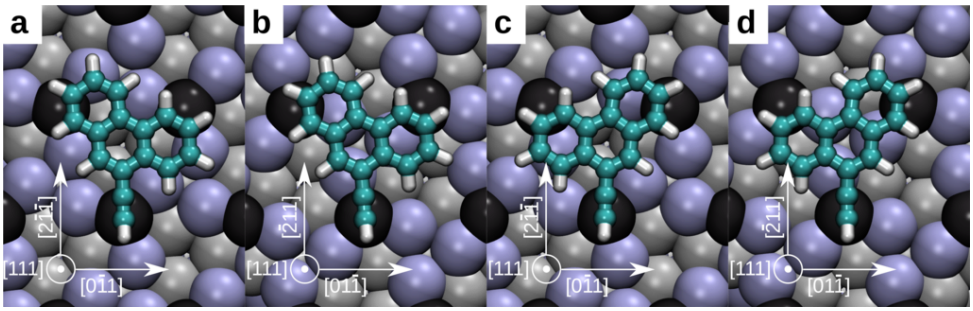


FIGURE 3.2: Representation of R and S adsorption geometries of 9-EP on top of the PdGa:A(111)Pd_1 surface. (a) R geometry, located aside of the Ga trimer; (b) R^* geometry, located on top of the Ga trimer; (c) S geometry, located aside of the Ga trimer. (d) S^* geometry, located on top of the Ga trimer. The gray (black at the surface) and violet spheres represent Pd and Ga atoms respectively. Cyan and white spheres represent C and H atoms of the 9-EP molecule, respectively.

All the proposed configurations were optimized with different DFAs and almost all of them (apart from BEEF-vdW [147]) predict R to be the most stable configuration. However, the energetic ordering of the remaining configurations (see Table 3.1) shows some discrepancies. This reveals the challenges in predicting enantioselectivity with state-of-the-art DFT approximations with different levels of vdW treatment.

The assignment of adsorption geometries to respectively R and S or R^* and S^* is better clarified by comparing simulated STM images to experimental ones (see Figure 3.1d,e; Figure 3.3). To elucidate the comparison, one should focus on the positions of the dark spots on the surface (whose position is on top of the subsurface Pd trimers). For R and S geometries the dark regions are located on the sides of the ethynyl "finger", while in the R^* and S^* the dark spots are located in front of it (for a comparison to the experimental images of the minority starred configurations we refer the reader to the original work by Prinz et al. [9]). The analysis of the

TABLE 3.1: Adsorption energies obtained with the different DFAs: adsorption energy/vdW contribution (adsorption height). All energies are in eV, all heights are in Å. Adsorption height is measured as the difference between the average positions of carbon atoms and Pd atoms of the first layer.

Fucntional	*R*	*R**	*S*	*S**
PBE [54]	-0.944/0.000 (2.46)	-0.901/0.000 (2.46)	-0.906/0.000 (2.50)	-0.922/0.000 (2.54)
PBE+DFT-D3 [54, 69]	-2.401/-1.543 (2.39)	-2.302/-1.497 (2.43)	-2.349/-1.537 (2.41)	-2.336/-1.523 (2.43)
revPBE+DFT-D3 [148, 69]	-2.589/-2.840 (2.33)	-2.490/-2.900 (2.33)	-2.560/-2.846 (2.34)	-2.481/-2.799 (2.37)
optB88-vdW [149]	-2.391/-3.166 (2.40)	-2.346/-3.099 (2.45)	-2.344/-3.158 (2.43)	-2.357/-3.090 (2.44)
rVV10-vdW [150]	-2.379/-1.930 (2.45)	-2.314/-1.885 (2.48)	-2.361/-1.901 (2.48)	-2.333/-1.933 (2.49)
BEEF-vdW [147]	-1.278/-2.090 (2.63)	-1.308/-2.132 (2.63)	-1.291/-2.028 (2.64)	-1.314/-2.050 (2.65)

details reveals matching of the geometries proposed in the experiment with the ones considered in the calculations.

The comparison of experimental and theoretical STM images (Figure 3.3) confirms that the geometries observed in the experiment correspond to *R* and *S* configurations.

3.2 Charge density difference in the adsorption of 9-EP

To understand the nature of 9-EP/PdGa:A(111)Pd_1 interaction in Figure 3.4a,b we show an isosurface of the charge density difference between the adsorbate-surface complex (9-EP@Pd_1) and the sum of the isolated systems (9-EP in gas phase and a clean Pd_1 surface). The plot reveals that the nature of 9-EP/Pd_1 interaction is not purely dispersive, a contribution from chemical bonding is also evident as a change of the hybridization state of the C-atom connected to H in the ethynyl group of 9-EP. Instead of being sp hybridized, as in the isolated molecule, the carbon atom becomes partially sp2 hybridized.

Such a transition can be noticed in Figure 3.4a,b where the red isosurface, showing a reduced charge density, is located parallel to the C-C triple bond and the green

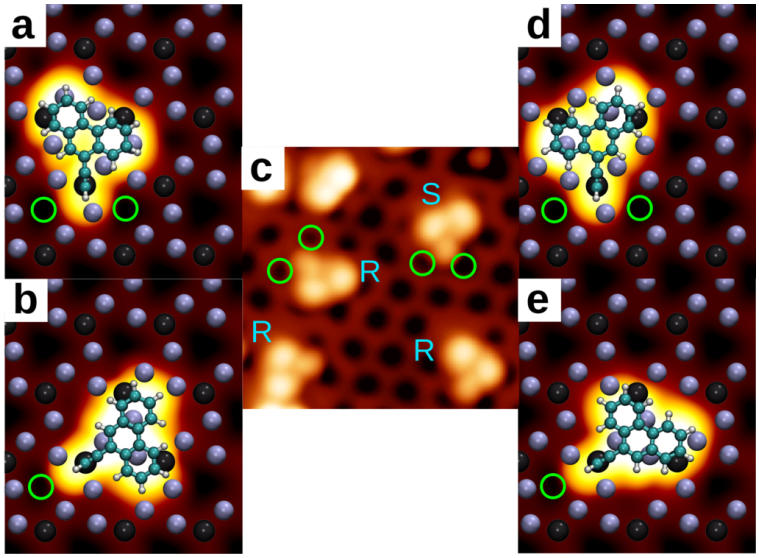


FIGURE 3.3: Comparison of experimental and simulated STM images of *R* and *S* enantiomers adsorbed on different adsorption sites. (a) Simulated image of *R* configuration. (b) Simulated image of R^* configuration. (c) Experimental STM image of *R* and *S* configurations. (d) Simulated image of *S* configuration. (e) Simulated image of S^* configuration. The green circles highlight dark spots on the surface to better recognize the orientation of the ethynyl "finger". *Experimental images courtesy of Samuel Stolz and Roland Widmer from the nanotech@surface laboratory at Empa.*

isosurface, showing an increased charge density, is located between both carbons and the Pd atom. Additionally, the geometrical changes confirm such a partial sp→sp2 transformation: hydrogen is pointing upwards with a C-C-H angle equal to 150° and the C-C bond length becomes 1.25 Å. This is an intermediate value for the typical double and triple bond lengths (1.20 Å and 1.34 Å respectively) between carbon atoms. We also note that the pattern of the charge density difference between ethynyl carbon and Pd looks very similar for the two enantiomers.

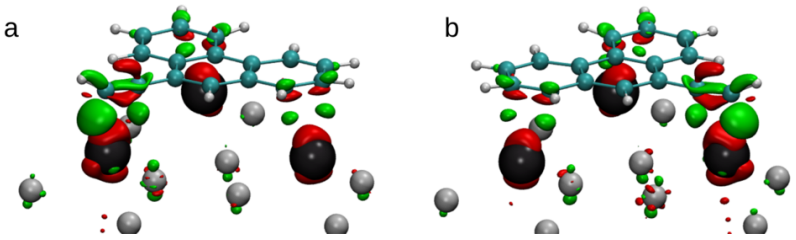


FIGURE 3.4: Charge density difference for (a) *R* and (b) *S* enantiomers revealing the partial chemical nature of the bonding between 9-EP and surface Pd atoms. The green/red isosurfaces show the excess/lack of charge density compared to the sum of the isolated 9-EP and clean PdGa:A(111)Pd_1 surfaces.

A plot of the plane average of the charge density difference along the surface normal (Figure 3.5) shows that most of the difference between the two configurations is located in the ethynyl group and in the region between the molecule and the surface.

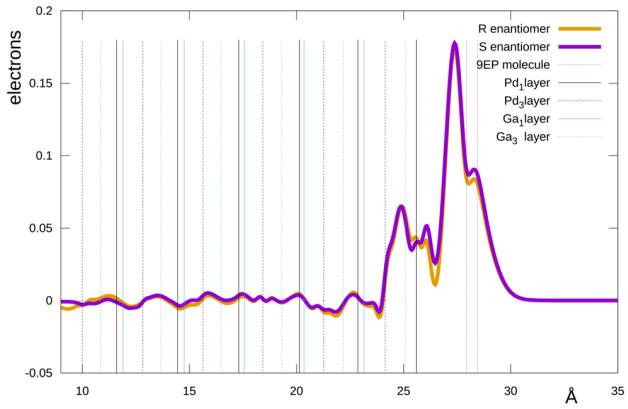


FIGURE 3.5: Charge density integral plotted as a function of the z coordinate. The integral is computed within each plane perpendicular to the [111] direction.

3.3 $R \leftrightarrow S$ transformation mechanism

Coexistence of dispersive interactions and chemical bonding results in a high adsorption energy, in the order of -2.5 eV, for 9-EP molecules on a Pd_1 surface. This finding is consistent with the experimental evidence that the molecules do not desorb from the substrate even after annealing cycles up to 670 K (temperature at which dimers are formed on the surface) and raises the question: how is it possible that adsorbed 9-EP molecules switch their chirality from S to R already at room temperature?

To answer this question, we enter the second analysis step with two hypotheses: either the ethynyl group can migrate via a surface mediated reaction from the 9^{th} to the 10^{th} carbon atom (while hydrogen does the opposite) or a low energy path exists for flipping of the molecule at the surface. We discarded our first hypothesis (as we will discuss below in more detail) since at a low temperature we could not observe any transition based on this mechanism: the molecule would rather break apart and its constituents chemically adsorb on the surface. On the other hand, the flipping

of 9-EP molecules requires overcoming a free energy barrier that we estimate in the order of 1.1 eV thus compatible with the experimental findings at room temperature.

To study the flipping mechanism we used metadynamics [113, 151, 112] within full DFT simulations. We defined the collective variable (CV) as the angle between the perpendicular to the plane of the 9-EP's central ring and the [111] direction (red arrow and white arrow in Figure 3.6a respectively).

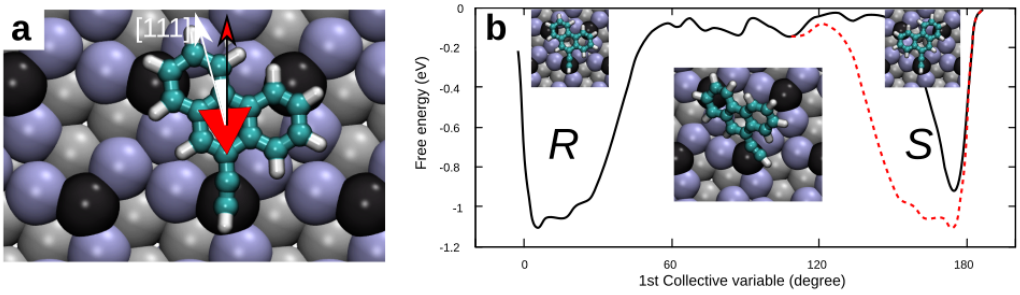


FIGURE 3.6: Flipping of the 9-EP molecule adsorbed on the PdGa:A(111)Pd_1 surface. (a) Description of the collective variable used in the metadynamics simulations to find a possible low energy path for the flipping of 9-EP adsorbed on the Pd_1 surface. (b) Approximate free energy landscape obtained after the first crossing has occurred; the barrier for the flipping is 1.1 eV (at 300 K). The red dashed line is a guide to the eye mimicking the second basin corresponding to the *S* configuration. Superimposed are atomistic sketches of the most relevant configurations for the 9-EP molecule along the flipping path.

Summing up all contributions to the history-dependent potential constructed during a converged metadynamics simulation allows to identify the free energy of the system along the set of CVs [151]. Our choice for the CV is rather straightforward and could result in an overestimation of the barrier separating the two free energy basins. Moreover, convergence of the metadynamics simulation is not feasible for such a system at the DFT level of theory. We stop our simulation after the system escapes for the first time the basin corresponding to the initial adsorption geometry (*R*) and starts filling the basin corresponding to the *S* adsorption geometry. It is reasonable to consider the depth of the basin obtained immediately after the first crossing as an upper limit for the free energy barrier of the flipping process [152]. In our case, for a simulation performed at 300 K, we obtain the mentioned barrier of 1.1 eV. Therefore, the flipping mechanism is compatible with the experimental evidence of a low temperature activation of the transition between the racemic mixture

and the *R* enantiomeric excess. In addition, we obtain an estimate of the potential-energy barrier for the flipping process (1.6 eV) performing a nudged elastic band (NEB) [119] simulation starting from the path identified along the metadynamics trajectory.

The computed barriers could be related to experimental conditions through the term $\Gamma e^{-\Delta G/k_B T}$ (ΔG being the free energy difference between the transition state and initial state) in the Arrhenius formula, if the attempt frequency Γ ere known. A direct calculation of Γ would be prohibitive, however assuming a value of 10^{13}, standard for processes at surfaces [116], reveals that below the dimerization temperature a flipping event happens on a timescale of seconds. Therefore, the flipping mechanism is compatible with the experimental evidence of low-temperature activation of the transition between the racemic mixture and the *R* enantiomeric excess.

We also explored a substrate mediated exchange of the H- and ethynyl groups. To do so, we employed metadynamics with two CVs: the distance between the 9^{th} carbon atom and ethynyl group as the CV1, and the distance between the H- group and 10^{th} carbon atom as the CV2, Figure 3.7. We also used multiple walkers metadynamics [113, 151, 112]. The simulations started with one walker in the *R* configuration and one in the *S* configuration.

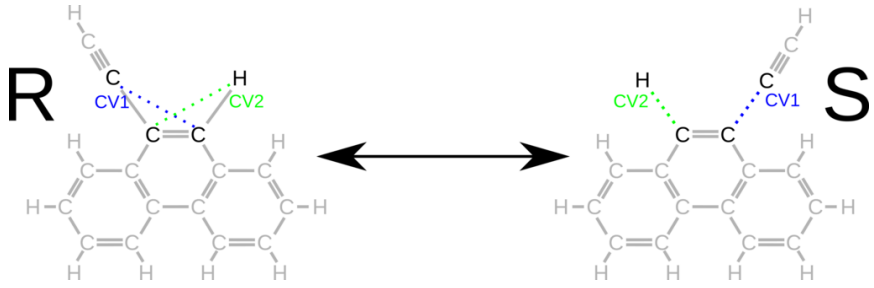


FIGURE 3.7: Description of the collective variables used in the metadynamics simulations to find a possible low energy path for the symmetric exchange of hydrogen and ethynyl group of 9-EP adsorbed on the Pd_1 surface.

A plot of the free energy, obtained after 2.5×10^5 MD steps, is shown in Figure 3.8. The exchange event did not occur, while for the molecule in the *S* configuration the hydrogen that initially belonged to the 10^{th} carbon detached from it and attached to the ethynyl group. This resulted in the appearance of the third cavity on the free energy surface. The depth of the *S* cavity remained $\approx$ 1.75 eV. Based on

these results, we concluded that according to the DFT the symmetric exchange of the H- and ethynyl group is significantly less favorable.

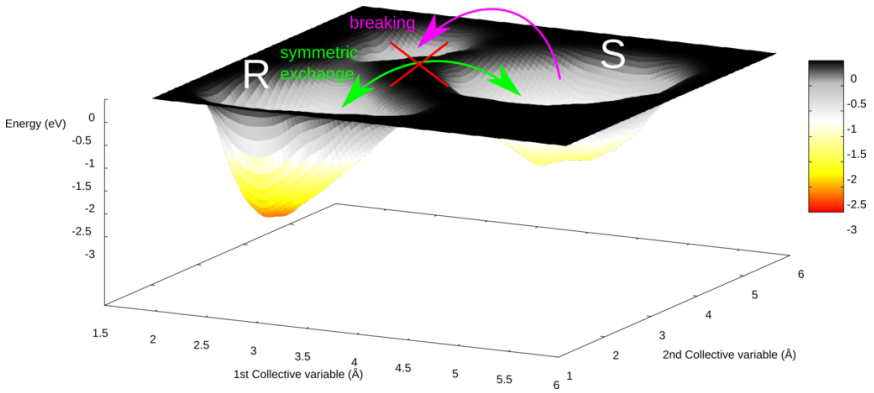


FIGURE 3.8: Free energy landscape obtained for the mechanism of a substrate mediated exchange of the H- and HCC- groups. The green arrow represents the mechanism that we aimed to simulate. The pink arrow shows the process that had actually occur: the hydrogen atom detached from the 10^{th} carbon atom of 9-EP and attached to the ethynyl group. Superimposed are atomistic sketches of the most relevant configurations for the 9-EP molecule along the flipping path.

3.4 Origin of the high enantioselectivity

At this stage, the third point of our investigation remains to be clarified: what is the origin of the high enantioselectivity experimentally observed? Given the impossibility to investigate properly the role of entropic/kinetic effects in the process, that would require at least properly converged metadynamics simulations, we explore in the following thermodynamic arguments based on the energy difference between R and S adsorption geometries ($E_{RS} = E_R - E_S$, where E_R and E_S are the computed adsorption energies of R and S respectively). In the experiments conducted by Prinz et al. [9], due to the high mobility of 9-EP on the Pd_1 surface there is no direct measurement of R/S enantiomeric excess at a specific temperature. To obtain a stable STM image the authors had to cool down the system to 5 K and, assuming that the cooling process was fast enough to freeze the equilibrium state obtained after annealing, they could identify the temperature dependence of the observed excess rate.

Under this hypothesis, we can estimate the free energy difference between the R and S configurations, referring to the information that at room temperature the enantiomeric excess for PdGa:A is 0.94 ± 0.02 [9]. The value $ee = 0.92$ (minimum

within the validity range) corresponds to an R to S ratio of

$$R/S = 1.92/0.08 = 24.00 \tag{3.1}$$

that is the equilibrium constant K_{eq} of the $S \rightleftarrows R$ transformation process. If the experimentally measured *ee* was obtained under equilibrium conditions at room temperature, the relationship $\Delta_r G^\circ = -kT \ln(K_{eq})$ would yield

$$\Delta_r G^\circ = -25.7 * \ln(24) = -81.7 \text{ meV} \approx -80 \text{ meV} \tag{3.2}$$

The Gibbs energy change at standard conditions $\Delta_r G^\circ$ can be expressed as follows

$$\Delta_r G^\circ = \Delta_r E_{tot} + \Delta_r F^\circ_{trans} + \Delta_r F^\circ_{rot} + \Delta_r F^\circ_{vib} + T\Delta_r S^0_{conf} + \Delta_r(PV) \tag{3.3}$$

where ΔF°_{trans}, ΔF°_{rot}, ΔF°_{vib} are the translational, rotational and vibrational free energy changes respectively. ΔS^0_{conf} is the configurational entropy change, P, V and T are the pressure, volume and temperature respectively. Since the $S \rightleftarrows R$ transformation happens on the surface with no volume change, we neglect the $\Delta_r(PV)$ term. The entropic term $T\Delta S^0_{conf}$ is also neglected, and we assume that the difference of translational, rotational or vibrational degrees of freedom between R and S is negligible so that

$$\Delta_r G^\circ \approx \Delta_r E_{tot} \tag{3.4}$$

which allows to compare $\Delta_r G^\circ \approx -80$ meV directly to DFT-computed total energy differences.

The chiral nature of the PdGa bulk can contribute to the 9-EP adsorption energy in two ways: through chemisorption contributions and through dispersion interactions.

We computed E_{RS} employing several vdW-inclusive DFAs. As previously discussed, at the beginning, we consider those available within the CP2K code distribution: PBE-D3 [54, 69], revPBE-D3 [148, 69], optB88-vdW [149], rVV10-vdW [150],

BEEF-vdW [147] as well as the conventional PBE [54] functional without vdW correction. In the PBE-D3 and revPBE-D3 calculations the C_9 term, describing three-body dispersion effects, was included and the damping function by Chai and Head-Gordon [153] was employed.

The values of E_{RS} are illustrated in Figure 3.9 decomposed into the vdW contribution and the remainder (abbreviated with *DFT* in the figure). E_{RS} varies considerably for the thus far mentioned methods, ranging from -0.052 eV to $+0.013$ eV, with BEEF-vdW being the only approach which describes the *S* configuration as the most stable one. All adsorption energies, as well as all corresponding adsorption heights, are present in Table 3.1. Figure 3.9 reveals that rVV10-vdW and BEEF-vdW imply a large vdW contribution to E_{RS}. However, in both cases the remaining DFT part is repulsive enough to reduce E_{RS} to less than 0.020 eV.

For all other DFAs, the DFT contribution amounts always to about -0.04 eV, while the vdW contribution is always smaller than 0.01 eV. Furthermore, the calculated adsorption heights differ up to 0.3 Å between all thus far discussed approaches. In general, the *R* enantiomer has a slightly smaller adsorption height than the *S* enantiomer. All calculations were performed with a surface model matching the PdGa lattice constant pertinent to the respective functional (see Table 3.1), yielding a spread of about 3 % among all discussed methods. PBE-D3 yields the smallest lattice constant and rVV10-vdW the largest. In order to determine if this lattice constant difference could lead to a qualitative change in the relative adsorption energy or the vdW contribution, we repeated the PBE-D3 calculation with the largest computed lattice constant. This approach is labeled PBE-D3* in Figure 3.9. In general, increasing the lattice constant leads to decreasing Pauli repulsion. This enables smaller adsorption heights which leads to a larger vdW attraction. Therefore, the absolute values of the adsorption energies increase in magnitude with an increasing lattice constant. However, the studied increase of the lattice constant results in an even positive vdW contribution to the adsorption energy difference E_{RS} and it also reduces the *DFT* contribution in magnitude, resulting in a decrease of E_{RS} by 0.02 eV. Increasing the lattice constant yields in this case ≈ 0.05 Å smaller adsorption heights

but results in no qualitative changes in the enantioselectivity, i.e. the vdW contribution in the D3 method has only a minor role in the adsorption energy difference. This is in contrast to the more accurate vdW methods as we will argue below.

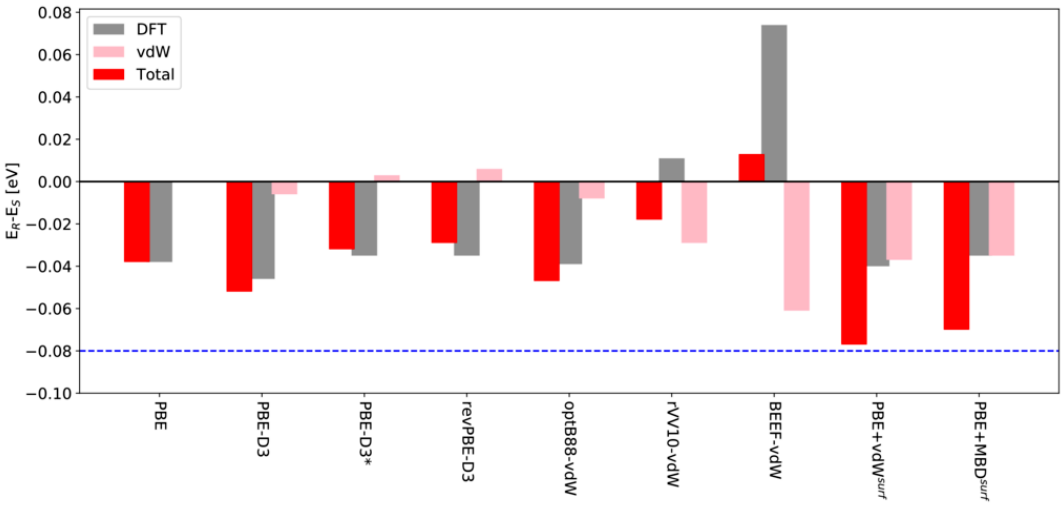


FIGURE 3.9: Adsorption energy differences E_{RS} obtained with several vdW-inclusive DFAs (total). The total adsorption energy difference is also decomposed into a dispersion contribution (vdW) and the remaining energy (DFT). The dashed blue line represents the minimal adsorption energy difference necessary to explain the experimentally observed enantiomeric excess. In the PBE-D3* approach, the lattice constant used for the generation of the surface was increased from 4.886 to 5.018 Å.

In general, although the thus far discussed results for E_{RS} indicate the R enantiomer as more stable than the S enantiomer, the adsorption energy differences are too small to explain the experimentally determined enantiomeric excess.

Therefore, we went one step further in the description of dispersive interactions employing the DFT+vdW^{surf} method [82], which accounts for the collective response of the underlying substrate by modeling screened vdW interactions based on the Lifshitz–Zaremba–Kohn theory [139, 83, 85, 84]. The vdW^{surf} is based on the pairwise Tkatchenko-Scheffler [74] (TS) dispersion model and includes screening effects of the substrate electrons by using renormalized "atom-in-a-solid" vdW parameters. The derivation of the parameters for Pd and Ga atoms within PdGa are discussed further in section 3.5.

The PBE+vdW^{surf} adsorption energies for both enantiomers were calculated using the full-electron code FHI-aims [92, 93, 94, 95]. In addition, we also computed the

adsorption energies with the many-body dispersion (MBD) method [76, 79] utilizing the renormalized vdW parameters from vdW^{surf}. This approach will be referred to as MBD^{surf}. The resulting adsorption energy differences are shown in Figure 3.9 and in Table 3.2.

The PBE+vdW^{surf} and PBE+MBD^{surf} approaches provide E_{RS} of -0.077 and -0.070 eV in favor of the R enantiomer, respectively. While the DFT contribution is comparable to the previous cases, the magnitude of the vdW contribution is now ≈ 0.04 eV, corresponding in both cases to about 50% of E_{RS}.

The results from both PBE+vdW^{surf} and PBE+MBD^{surf} are compatible with the experimental finding, confirm the thermodynamic origin of the enantioselectivity and highlight the dominant role of dispersion interactions in the enantioselectivity. We stress that only the PBE+vdW^{surf} and PBE+MBD^{surf} methods yield E_{RS} values large enough to correspond to the experimentally observed enantiomeric excess of the R enantiomer.

We now discuss the origin of the observed vdW-stabilization of the R enantiomer in the PBE+vdW^{surf} method. Figure 3.10a,b shows histograms of the absolute value of the vdW interaction between the adsorbed molecule and the surface atoms at a given interatomic distance for both enantiomers (E^{vdW}). The main qualitative difference between the two enantiomers originates from adsorbate-Ga interactions at small distances. In Figure 3.10c we show the vdW interaction energy difference between the R and S enantiomers when only contributions between atoms having interatomic distances smaller than a cutoff distance r are included. It can be seen that stabilization of the R enantiomer with respect to the S enantiomer mainly originates from distances smaller than 5 Å (2.5-5.0 Å range). Note that the difference in the interaction energies is not exactly equal to the difference in the adsorption energies, since this analysis does not account for differences originating from within the adsorbate and from within the metal surface.

To better illustrate the geometrical dissimilarity between R and S enantiomers, we plotted the number of C-Ga pairs as a function of interatomic distance (Figure 3.10f). The plot is aligned to the graph of Figure 3.10c. The subplot in Figure 3.10f focuses on the region 2.5-5.0 Å. The first relevant vdW contribution comes

TABLE 3.2: Adsorption energies of *R* and S configurations obtained with the different vdW-inclusive DFAs: adsorption energy/vdW contribution (adsorption height). All energies are in eV, all heights are in Å. Adsorption heights are measured as the difference between the average positions of carbon atoms and Pd atoms of the first layer. The adsorption energy difference, together with the vdW contribution to that difference, is given in the fourth column. The PdGa lattice constant value in Å is given in the fifth column.

DFA	*R* enantiomer	*S* enantiomer	$E_R - E_S$ [eV]	Lattice constant [Å]
PBE	-0.944/0.000 (2.46)	-0.906/0.000 (2.50)	-0.038/0.000	4.971
PBE-D3	-2.401/-1.543 (2.39)	-2.349/-1.537 (2.41)	-0.052/-0.006	4.886
PBE-D3[a]	-2.452/-1.522 (2.34)	-2.420/-1.525 (2.34)	-0.032/0.003	5.018
revPBE-D3	-2.589/-2.840 (2.33)	-2.560/-2.846 (2.34)	-0.029/0.006	4.891
optB88-vdW	-2.391/-3.166 (2.40)	-2.344/-3.158 (2.43)	-0.047/-0.008	4.974
rVV10-vdW	-2.379/-1.930 (2.45)	-2.361/-1.901 (2.48)	-0.018/-0.029	5.018
BEEF-vdW	-1.278/-2.090 (2.63)	-1.291/-2.028 (2.64)	0.013/-0.061	5.012
PBE+vdW^{surf} [83]	-2.609/-1.883 (2.14)	-2.532/-1.845 (2.16)	-0.077/-0.037	4.959
PBE+MBD^{surf} [76, 79]	-2.354/-1.629 (2.24)	-2.284/-1.594 (2.26)	-0.070/-0.035	4.895

[a]The lattice constant for this calculation was set to the largest DFT value, obtained with the rVV10-vdW functional: 5.018 Å.

from the C-Ga pairs whose interatomic distances are smaller than 3 Å. Those interactions favor the *R* configuration with respect to *S*, since *S* does not have any C-Ga contributions at this range (see Figure 3.10b,f). Then every half of angstrom we observe alternating results with respect to *R* and *S* stability: 3.0 – 3.5 (*R* is the most stable), 3.5 – 4.0 (*S*), 4.0 – 4.5 (*R*), 4.5 – 5.0 (*S*). At distances larger than 5.0 Å E_{RS}^{vdW} essentially becomes a straight line with minor oscillations.

In order to demonstrate the importance of dielectric screening, Figure 3.10c-d contain the same analysis performed using the PBE+vdW^{surf} structures but evaluating the vdW energies with the traditional Tkatchenko-Scheffler (TS) approach [74],

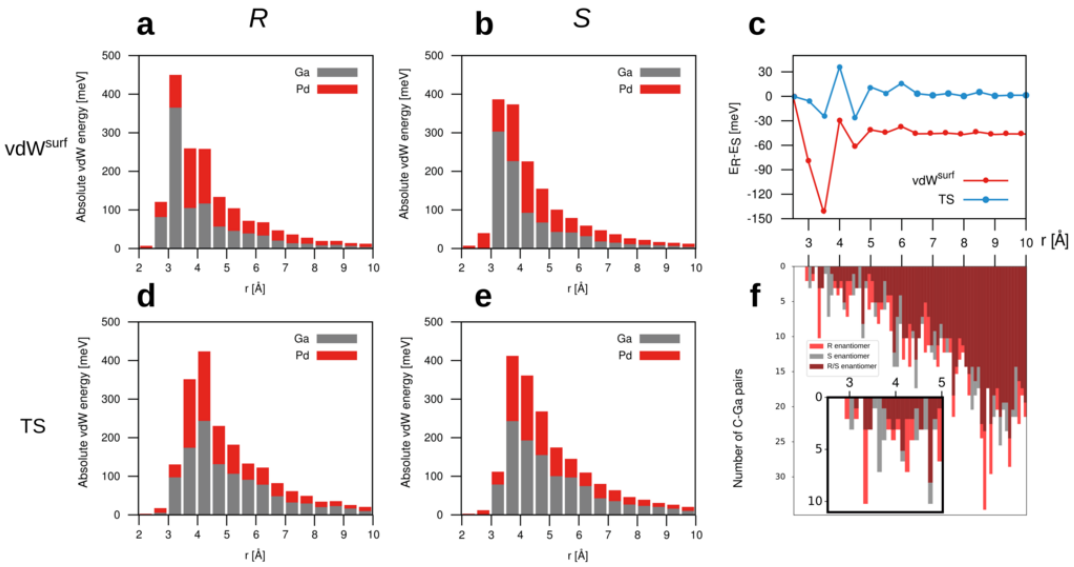


FIGURE 3.10: (a,b,d,e) Histograms of the absolute value of the vdW interaction between the adsorbed molecule and the surface atoms for a given interatomic distance: vdW^{surf} interaction energies for the *R* (a) and *S* (b) enantiomer; vdW interaction energies obtained with the traditional Tkatchenko-Scheffler (TS) approach without screened free-atom parameters but calculated with the PBE+vdW^{surf} geometries for the *R* (d) and *S* (e) enantiomer, respectively. The values are decomposed in contributions between adsorbate atoms and Pd and Ga atoms, respectively. (c) E_{RS}^{vdW} — difference in vdW interaction energy between the adsorbate molecule and the surface atoms contributed by interactions between atoms which have an interatomic distance smaller than *r*. (f) Number of C-Ga pairs for a given atomic distance (aligned to the plot of c).

i.e. using free-atom parameters as a starting point. In this case, there is almost no difference between the two enantiomers (as for the case of DFT-D3 results).

Accounting for dielectric screening decreases the magnitude of the adsorption energies and modifies the range of the employed damping function. Here, this leads to an increased stabilization of the *R* enantiomer originating from short-range adsorbate-Ga interactions. Indeed, the most important effective change in the interaction scheme by moving from TS to vdW^{surf} is the rescaling of the vdW radius R_{vdW} from the TS value of 2.22 Å to 1.30 Å. We note that this change directly originates from the dielectric constant determined from first-principles calculations. This suggests that the screening properties of the substrate material strongly affect the appropriate weighting of the interactions at intermediate distances, which in turn is the

factor determining enantioselectivity. Our analysis also demonstrates that the gallium atoms in PdGa do not simply act as “spacers": they actively affect the adsorption energies of carbon-based molecules, as was recently pointed out by Bechthold et al. [154].

3.5 Methods

All the simulations, whose results are shown in the Table 3.1, as well as all the metadynamics simulations were performed by means of the CP2K code [96, 97] implementing DFT within a mixed Gaussian plane waves approach [98]. The surface/adsorbate systems were modeled within the periodically repeated slab scheme [155] and the Martyna-Tuckerman approach [87] was employed to decouple the surfaces from their replica. The electronic states were expanded with the TZV2P Gaussian basis set [156] for C and H atoms and the DZVP basis set for Pd and Ga. A cutoff of 1200 Ry was used for the plane waves basis set. Goedecker-Teter-Hutter pseudopotentials [157] were used to represent the frozen core of the atoms. Eighteen and thirteen electrons were included in the valence for the Pd and Ga species respectively.

For all the simulations (except metadynamics) the model system contained 23 atomic layers of PdGa along the [111] direction, the adsorbate and 40 Å of vacuum. As indicated in Table 3.1, we used different possible exchange correlation functionals and parameterizations for the van der Waals interactions. For each functional we determined the corresponding bulk lattice parameter of PdGa. To optimize the adsorption geometries, we considered supercells of 20.76x23.96 Å corresponding to twelve surface units. We kept the atomic positions of the bottom 6 layers of the slab fixed to the ideal bulk positions, all other atoms were relaxed till forces were lower than 0.005 eV/Å. The experimental results presented by Prinz and co-workers did not highlight any coverage dependent effect. Both low coverage and high coverage data are available. The supercell adopted in the simulations corresponds to a distance between an adsorbed molecule and its replica of 12.5 Å, which can be considered corresponding to a “low coverage” situation. To obtain simulated STM

images [126, 127], within the Tersoff-Hamann approximation [158, 125] we extrapolated the electronic orbitals to the vacuum region following the approach outlined by Tersoff [125].

For metadynamics we considered 7 atomic layers (approx. 5 Å) along the [111] direction, the adsorbate and 20 Å of vacuum. The supercell size was 13.83x23.96 Å corresponding to eight surface units. We kept the atomic positions of the bottom 3 layers of the slab fixed to the ideal bulk positions. This allowed us to perform trajectories of approximately 15 ps and 125 ps respectively for the flipping and symmetric exchange processes.

All calculations within the all-electron code FHI-aims [92, 93, 94, 95] were performed by using the Perdew-Burke-Ernzerhof (PBE) density functional [54]. The vdW interactions have been calculated with the vdW$^{\text{surf}}$ [82, 83, 139] and the many-body disperision (MBD) method [76, 79]. vdW$^{\text{surf}}$ is based on the pairwise Tkatchenko-Scheffler [74] approach but includes screening effects of the substrate electrons by using re-scaled free-atom vdW parameters. The MBD method has been used with the re-scaled parameters from vdW$^{\text{surf}}$ and will be referred to as MBD$^{\text{surf}}$. In order to apply these methods, we have derived the screened vdW parameters for a Pd and Ga atom within PdGa. Since no viable experimental measurement was available, the dielectric function was calculated within FHI-aims by using the random-phase approximation utilizing the PBE functional and tight species default settings [159]. The real part of the dielectric function ϵ_1 on the imaginary frequency axis (iu) was obtained via the Kramers-Kronig relation

$$\epsilon_1(iu) = 1 + \frac{2}{\pi}\int_0^\infty \frac{\epsilon_2(\omega)}{u^2+\omega^2}\, d\omega \tag{3.5}$$

with $\epsilon_2(\omega)$ being the imaginary part of the dielectric function on the real frequency axis. The screened parameters can now be derived as described in the references [82, 83]. In order to assess the quality of our calculated dielectric function, we first derived the parameters for bulk Pd and compared them with the values given in the reference [82] before moving on to PdGa (see Table 3.3). The obtained parameters agree very well (within 7%) with the reference values, which have been determined based on experimental measurements.

TABLE 3.3: Screened dispersion parameters for atoms in the bulk: polarizabilities (α), C_6 coefficients, and vdW radii (R_{vdW}).

Atom	α [bohr^3]	C_6 [hartree$\times$bohr^6]	R_{vdW} [bohr]
Pd (ref) [82]	13.9	102	3.06
Pd (calc)	13.4	95	3.03
Pd (in PdGa)	13.0	103	3.00
Ga (in PdGa)	11.9	87	2.45

For PdGa we used the same approach but calculated screened vdW parameters corresponding to a unit of one Pd and one Ga atom, which amount to a polarizability $\widetilde{\alpha} = 24.9$ bohr^3 and a dispersion coefficient $\widetilde{C_6} = 378$ hartree$\times$bohr^6. In order to extract separate parameters for Pd and Ga we essentially followed an approach suggested in the reference [160]. We assume that the polarizabilities are additive and approximate $\widetilde{\alpha}$ by

$$\widetilde{\alpha} \approx \alpha^{\mathrm{Pd}} + \alpha^{\mathrm{Ga}} \tag{3.6}$$

The screened atomic polarizabilities can now be obtained as [160]

$$\alpha^{\mathrm{Pd}} = \frac{\widetilde{\alpha}\kappa}{\kappa + 1} \tag{3.7}$$

$$\alpha^{\mathrm{Ga}} = \frac{\widetilde{\alpha}}{\kappa + 1} \tag{3.8}$$

with $\kappa = V^{\mathrm{Pd}}/V^{\mathrm{Ga}} = 1.088$ being the ratio of the Hirshfeld volumes (V) in the solid. Furthermore, we express $\widetilde{C_6}$ as the sum of all homonuclear and heteronuclear dispersion coefficients [82, 160]

$$\widetilde{C_6} = C_6^{\mathrm{Pd}} + C_6^{\mathrm{Ga}} + \frac{4C_6^{\mathrm{Pd}}C_6^{\mathrm{Ga}}}{\frac{\alpha_{\mathrm{Ga}}}{\alpha_{\mathrm{Pd}}}C_6^{\mathrm{Pd}} + \frac{\alpha_{\mathrm{Pd}}}{\alpha_{\mathrm{Ga}}}C_6^{\mathrm{Ga}}} \tag{3.9}$$

By using the relation $C_6^{\mathrm{Pd}} = \kappa^2 C_6^{\mathrm{Ga}}$ we can now calculate our screened dispersion coefficients. Finally, the vdW radii R_{vdW} are calculated for atom i by

$$R_{\mathrm{vdW}}^{i} = \left(\frac{\alpha^{i}}{\alpha_{\mathrm{free}}^{i}}\right)^{1/3} R_{\mathrm{vdW}}^{\mathrm{i,free}} \tag{3.10}$$

where α^{i}_{free} and $R^{\text{i,free}}_{\text{vdW}}$ are the polarizability and vdW radius for the free atom i, respectively [74]. The so obtained screened vdW parameters used in the vdW$^{\text{surf}}$ and MBD$^{\text{surf}}$ approach for PdGa are shown in Table 3.3.

The unit cell of PdGa was optimized with PBE+vdW$^{\text{surf}}$ and PBE+MBD$^{\text{surf}}$ within FHI-aims using a 12×12×12 k-grid and tight species default settings. The obtained lattice parameters are shown in Table 3.4 and have been used to create the surface structures for each method. It can be seen that the lattice constant obtained with PBE+MBD$^{\text{surf}}$ agrees almost perfectly with the experimental value and PBE+vdW$^{\text{surf}}$ slightly overestimated the lattice constant.

TABLE 3.4: Optimized lattice constants (a) of PdGa.

Method	a [Å]
Exp. [7]	4.897
PBE+vdW$^{\text{surf}}$	4.959
PBE+MBD	4.911
PBE+MBD$^{\text{surf}}$	4.895

The adsorbate/surface complex was optimized for both enantiomers with all previously mentioned methods within FHI-aims. All calculations were performed with a 2×2×1 k-grid, tight species default settings for integration grids and basis functions, and the scalar-relativistic ZORA (zero-order regular approximation) approach. Convergence criteria of 10^{-5} eV and 10^{-5} electrons/Å^3 were used for the total energy and the charge density, respectively. The structures have been optimized until all atomic force components were smaller than 10^{-2} eV/Å.

3.6 Conclusions

In summary, we provided evidence of the van der Waals nature of the enantioselectivity properties of the PdGa:A(111)Pd$_1$ surface with respect to the adsorption of 9-ethynylphenanthrene. The chirality of the flat surface originates from the chiral nature of the PdGa crystal and expression of enantioselectivity was first reported by Prinz et al. but the chemical and physical origin of it was unknown up to now. Such investigation reveals the possibility of a van der Waals driven chiral recognition on a solid chiral surface.

After demonstrating the correctness of the adsorption geometries proposed in the experiments, we proved that flipping of the molecules on the surface is possible below dimerization temperature a and is indicated as the mechanism for the transition from a racemic mixture to enantiomeric excess. Finally, we analyzed in detail the molecule/substrate interaction and identified the van der Waals contribution from the subsurface Ga atoms as the main factor in the enantioselectivity. Our work, based on recently developed van der Waals-inclusive density-functional approximations, which include dielectric screening effects, stresses the importance of proper modeling of dispersion interactions in the description of adsorbate–substrate systems. Indeed, simplified models for the dispersion interactions without account for dielectric screening effects did not allow to explain the enantioselectivity of the PdGa surface. In contrast, a proper description of dielectric screening leads to an appropriate weighting of the vdW interactions at intermediate distances (2.5 - 5 Å) and to an energy landscape in agreement with the experiments.

Chapter 4

Adsorption of 9-EP molecules on PdGa:A($\bar{1}\bar{1}\bar{1}$) surface

This chapter is dedicated to rationalizing the experimental and simulated data available for the adsorption of a 9-EP molecule on the PdGa:A($\bar{1}\bar{1}\bar{1}$)Pd_3 surface. As in the case of Pd_1 discussed above, also here experiments reveal chirality-induced motifs for the surface patterns. The topmost layer of PdGa:A($\bar{1}\bar{1}\bar{1}$)Pd_3 (further referred as Pd_3) differs from the Pd_1 (see section 1.3) as the Pd isolated sites are replaced by isolated Pd trimers [12]. The enantioselective properties of this surface are not yet fully understood. The deposition of 9-EP molecules at room temperature on the Pd_3 is not accompanied by enantioselectivity; however, further annealing to 400 K or higher temperatures induces the formation of 9-EP trimers and, despite isolated molecules being a racemic mixture, the majority of trimers have a well-defined chirality. The trimers appear to be stable upon annealing, revealing that intermolecular interactions within a trimer are not purely dispersive. Inference from STM images leads to the investigation of the presence of an atom acting as a ligand center in the middle of the trimer.

In this chapter we analyze various hypotheses. We show that a single Pd atom is not stable in the center of three Pd atoms where the 9-EP molecules bind to each other, thus it cannot stabilize a 9-EP trimer. According to our investigations, the CO molecule, immensely present on the surface, could also not stabilize it by just being in the center. The chemical reaction between 9-EP and CO resulting in the

formation of tertiary alcohol could potentially explain the stabilization but was disproved based on the results of the nc-AFM experiment. Only the hypothesis with the Ga-coordinated trimer remained valid, but its applicability depends on the initial hydrogenation state of the 9-EP molecule.

4.1 9-EP trimers adsorbed on a Pd_3 surface

According to the experimental results, the majority ($\sim$ 99%) of trimers formed on a Pd_3 surface upon annealing to 490 K are composed of 9-EP molecules in the *R* configuration (Figure 4.1). The trimers are stable up to 570 K, while annealing up to 670 K results in irreversible chemical transformations and the disappearance of the trimers.

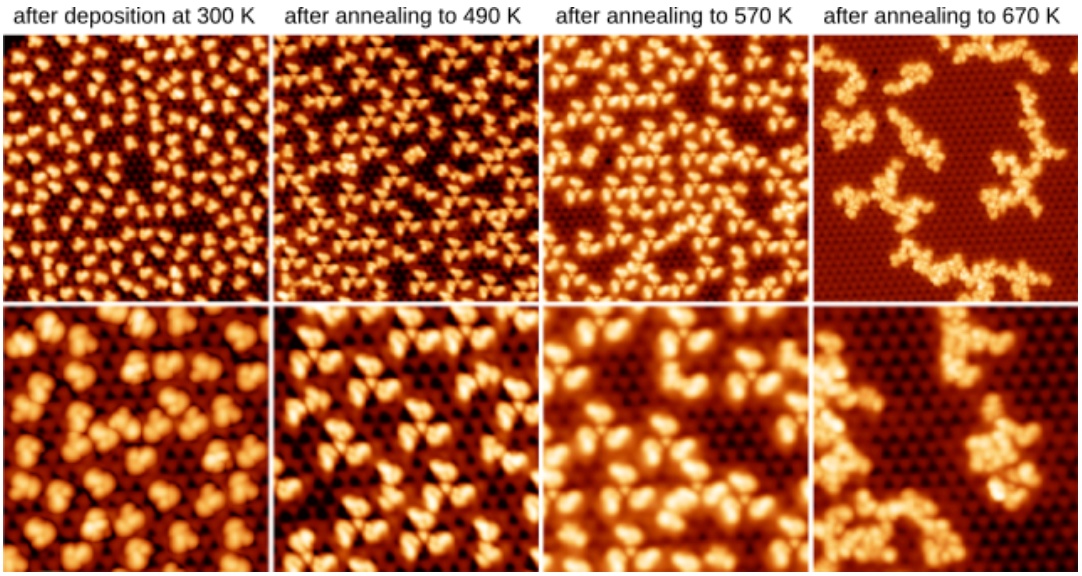


FIGURE 4.1: Experimental STM images of 9-EP molecules adsorbed on a Pd_3 surface. The upper panels correspond to an area of 20×20 nm, while the lower images span 10×10 nm. The deposition was performed at room temperature and the system was further annealed for 80 minutes at 490, 570 and 670 K. After the deposition and annealing steps, STM measurements were taken at 5 K (bias voltage 20 mV). Annealing to 670 K activated irreversible chemical transformations of the 9-EP molecules and no trimers are left. *Experimental images courtesy of Jan Prinz, Roland Widmer and Oliver Gröning from the nanotech@surface laboratory at Empa.*

As for the case of the Pd_1 surface, the starting point was to determine the adsorption geometries of monomers and trimers by comparing experimental and simulated

STM images. Figure 4.2a displays a detail of isolated 9-EP molecules in *R* and *S* configurations and Figure 4.2b shows an overlap of an experimental STM image with ball-stick models of 9-EP. The experimental STM images match the simulated ones, validating the proposed adsorption geometries for both enantiomers.

Curiously, looking carefully at the adsorption geometries one could notice a difference in the appearance of *R* and *S* configurations Figure 4.2b. The "ethynyl finger" (pointed out on the image with white arrows) looks less pronounced for *R* compared to *S*. This could suggest hydrogen loss in the ethynyl group already at room temperature. To clarify this point, we simulated STM images of both 9-EP configurations with and without the extremal hydrogen atom (Figure 4.2).

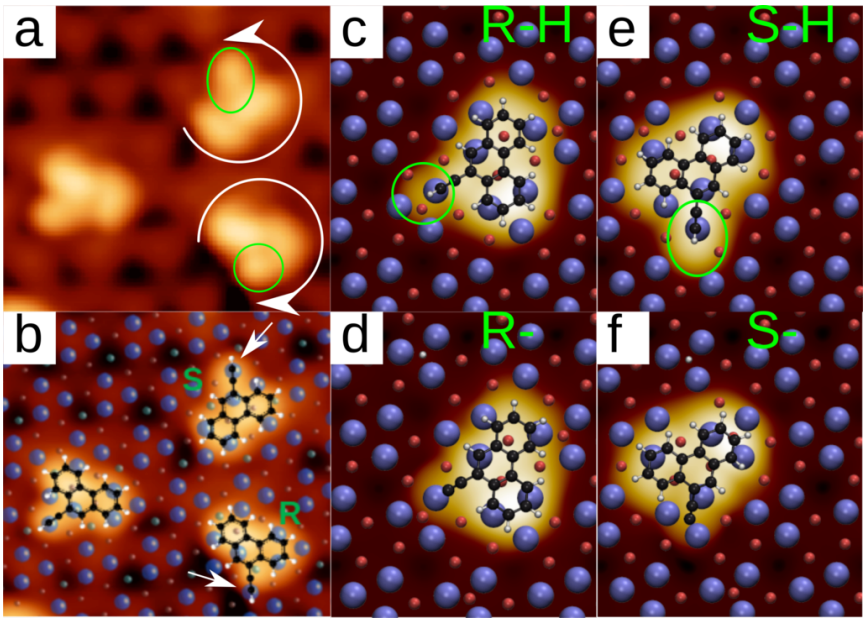


FIGURE 4.2: (a) Zoom of high-resolution STM images of a 9-EP molecule on the Pd_3 surface, highlighting the difference between *R* and *S* enantiomers with white arrows and green circles. (b) Overlap of experimental STM images with a ball-stick model of a 9-EP molecule. Simulated STM image of an *R* configuration 9-EP molecule with (c) and without (d) hydrogen attached to the ethynyl group. Notice the significant difference between the ethynyl fingers for the hydrogenated *R* and *S* configurations. *Experimental images courtesy of Samuel Stolz, Roland Widmer and Oliver Gröning from the nanotech@surface laboratory at Empa.*

Based on the obtained results, we tend to exclude the hypothesis of dehydrogenation at room temperature for several reasons. Looking at the simulated STM in Figure 4.2c,e it is clear that the appearance of "ethynyl fingers" in both hydrogenated

R and *S* configurations is substantially different. Indeed, in the *R* case it appears significantly shorter and less intense. Another reason is that STM images obtained at 300 K right after the deposition reveal that the 9-EP monomers are mobile on the surface. DFT simulations imply that the dehydrogenated ethynyl group binds to the surface Pd atom, thus hindering the mobility of the radical 9-EP compared to the intact molecule. The binding of the radical to the Pd trimer is strong and can be roughly estimated computing the adsorption energy for both *R* and *S* configurations in the hydrogenated (*R*, *S*) and dehydrogenated (*R*−, *S*−) case: −5.00 and −3.01 eV for *R*− and *R* respectively, −5.09 and −2.96 for *S*− and *S* respectively.

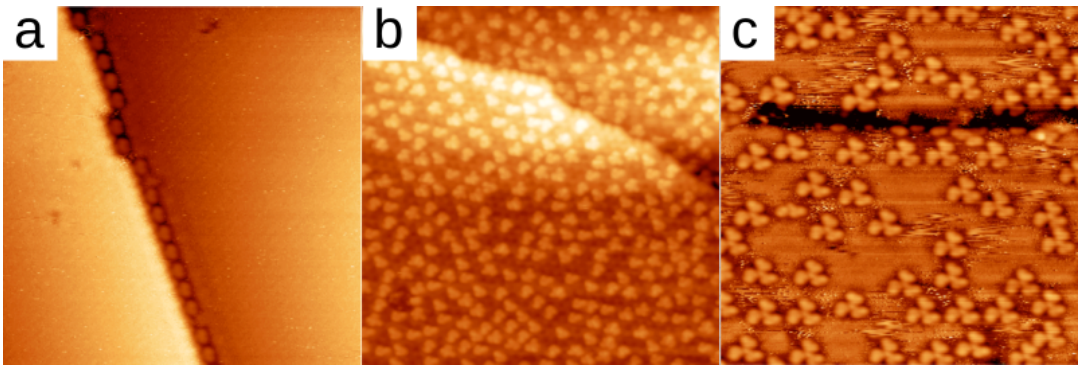


FIGURE 4.3: (a) An experimental STM image of 9-EP adsorbed on Pd_3 obtained at room temperature. (b) Same as (a) but acquired at temperature of liquid nitrogen. (c) An STM image obtained at room temperature after anneling to 550 K. *Experimental images courtesy of Samuel Stolz, Roland Widmer and Oliver Gröning from the nanotech@surface laboratory at Empa.*

The situation changes when the system is heated to more then 400 K. At elevated temperatures, the 9-EP molecules (apart from forming trimers) also form dimers. The structure of a dimer was established with the help of the nc-AFM technique. To form such a structure, the 9-EP molecules most probably lose hydrogen atoms from the ethynyl groups and then the radicals can bind to each other. Similar behavior was also observed on the Pd_1 surface, where molecules started to dimerize also at 400 K.

Therefore, we can conclude that at temperatures higher than 400 K, the 9-EP molecules most probably lose hydrogen atoms of the ethynyl group and become unstable. This reactive state is assumed to be the starting point for further dimers and trimers formation. Moving to the analysis of trimers, we again compare the simulated and experimental STM images Figure 4.5.

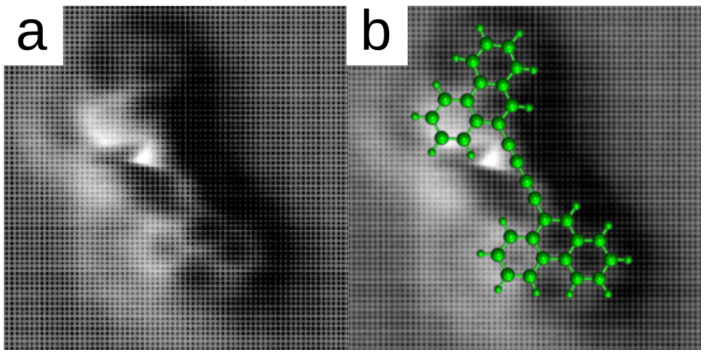


FIGURE 4.4: (a) A nc-AFM image of 9-EP dimer formed on the surface at 400 K. (b) Overlapping of a nc-AFM image of 9-EP molecule with its chemical structure reveals that dimers are formed from two radicals of 9-EP molecules. *Experimental image courtesy of Samuel Stolz, Roland Widmer and Oliver Gröning from the nanotech@surface laboratory at Empa.*

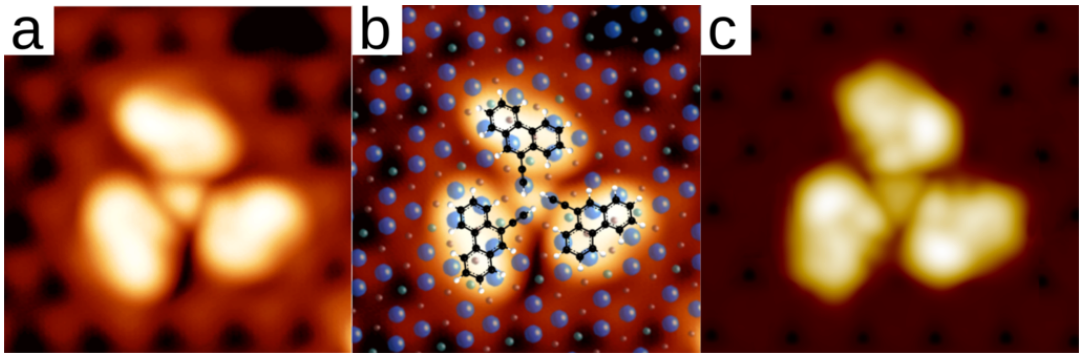
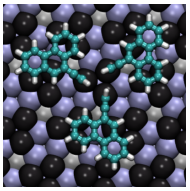
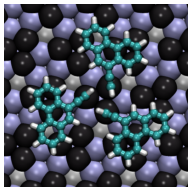
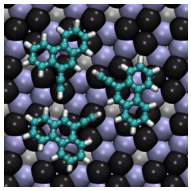
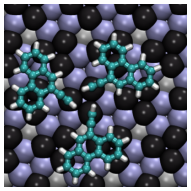
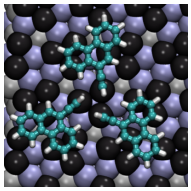


FIGURE 4.5: (a) Details of an isolated 9-EP trimer on the Pd_3 surface. (b) Overlap of the experimental image and a ball-stick model of a 9-EP molecule. (c) A simulated STM image matching the experiment. *Experimental image courtesy of Samuel Stolz, Roland Widmer and Oliver Gröning from the nanotech@surface laboratory at Empa.*

In the trimer, the geometry of 9-EP is very similar to the geometry of an isolated molecule in the *R* configuration. Assuming for the beginning that 9-EP is not undergoing chemical reactions upon formation of the trimers, we computed the energies of several possible configurations formed starting from 9-EP molecules on neighboring Pd_3 sites as shown in Table 4.1.

Being clear that in terms of energetics a trimer composed by molecules in *R* configurations is lower in energy compared to all other possible *R* and *S* configurations, we focus now on understanding which interactions stabilize the trimers and what is the origin of the bright protrusion at the center of the trimer (see Figure 4.5). We analyze six hypotheses presented in Figure 4.6. We examined them based on geometric, energetic and spectroscopic arguments. Hypothesis (1) was discarded because of unstable geometry: the 9-EP molecules would adopt a configuration too compact

TABLE 4.1: Trimer's adsorption energy (eV) per molecule obtained for different possible configurations of the 9-EP molecules

R_1	R_2	S_1	S_2	S_3
				
Adsorption energy per molecule				
−3.09	−2.84	−2.91	−2.78	−2.92
Energy difference with respect to the most stable conformation (per monomer)				
0.0	0.25	0.18	0.31	0.17
Energy difference with respect to the most stable conformation (per trimer)				
0.0	0.75	0.54	0.93	0.51

compared to the features appearing in the experimental STM images. Moreover, forming a benzene ring out of ethynyl groups of the 9-EP molecules would cause steric hindrance: the phenanthrene part of the molecule would tilt rather then staying planar, as it was observed experimentally.

Following the proposed hypotheses, we estimate the trimer formation energy needed to reproduce the concentration of trimers and dimers obtained in the experiments considering dimer and trimer formation energies as parameters in a series of Monte Carlo simulations.

4.2 Monte Carlo study of trimers formation

In this study we performed Monte Carlo (MC) simulations based on the Metropolis algorithm [161]. We compared equlibrium distributions to the data available from STM images obtained at low coverage that are summarized in Figure 4.7.

The trimer formation energy is derived assuming that experimental STM images at low coverage correspond to a thermodynamic equilibrium. The molecules were represented through simplified shapes that mimic the size of 9-EP and reproduce correctly the occupation of Pd sites (Figure 4.8). 1200 molecules were randomly positioned on a model Pd_3 surface measuring $\sim$ 100nm $\times$100nm. The occupied and unoccupied sites were distinguished in such a way as to respect the preferred

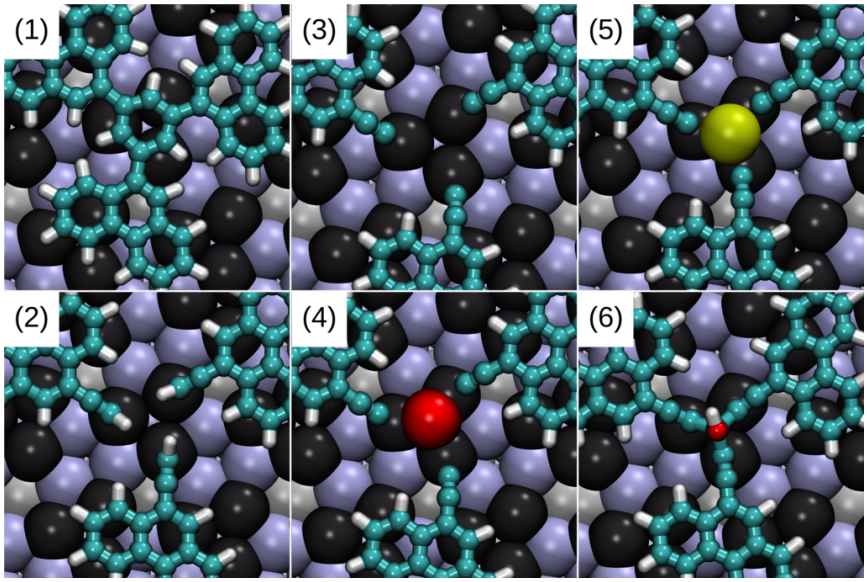


FIGURE 4.6: Hypotheses proposed to explain the stabilization of the 9-EP trimers. (1) The ethynyl groups trimerise and form a benzene ring. (2) The trimer is formed solely by three intact 9-EP molecules. (3) The trimer is formed by three 9-EP molecules with dehydrogenated ethynyl groups. (4) Three 9-EP molecules form covalent bonds through the ethynyl group pointing towards the central Pd atom located in the Pd trimer. (5) Three 9-EP molecules form covalent bonds through the ethynyl group pointing towards the central Ga atom located in the Pd trimer. (6) Three 9-EP molecules form covalent bonds with the CO molecule in the center to form a tertiary alcohol.

adsorption geometry of 9-EP (Figure 4.8a) as it is shown in Figure 4.8d where the free sites around a single "molecule" are marked with green circles.

The energy of a configuration in the Monte Carlo trajectory was defined by the number of trimers (only trimers composed by R molecules were considered) and by the number of dimers (dimers corresponded to a trimer missing one molecule)

$$E_{tot} = N_d E_d + N_t E_t \tag{4.1}$$

where N_t/N_d is the number of dimers/trimers and E_d/E_t is the formation energy of dimers/trimers. A step of the MC simulation consisted of a random selection of a molecule, moving the selected molecule to a random position, accepting the move with probability $p = \exp(-\Delta E/(k_B T))$ in the case of an energy increase and with probability $p = 1$ in the case of an energy decrease. ΔE is the energy change, k_B is the Boltzmann constant, T is the temperature (set to 600 K for all the simulations).

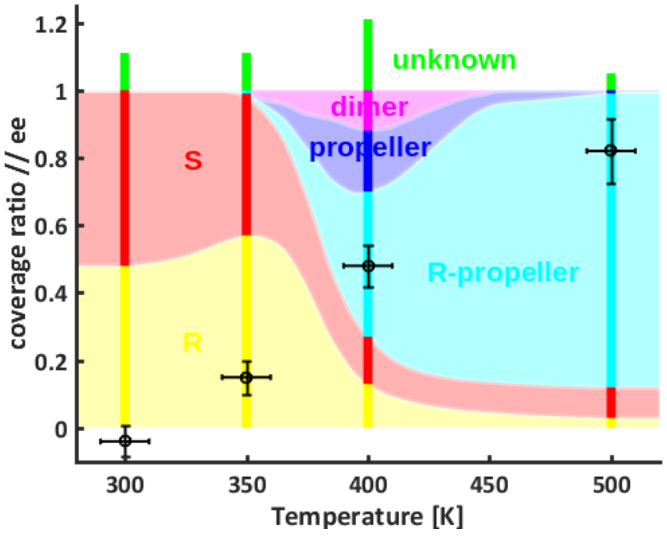


FIGURE 4.7: Experimental results based on the STM data obtained at different temperatures at low coverage. *Experimental image courtesy of Samuel Stolz, Roland Widmer and Oliver Gröning from the nanotech@surface laboratory at Empa.*

Long[1] MC simulations were performed for each couple of E_d and E_t chosen on a grid limited by $E_d = [-0.005 : -0.01]$ and $E_t = [-0.05 : -1.0]$.

In Figure 4.9 the number of trimers at equilibrium is plotted as a function of E_d and E_t. This plot reveals that to have the same percentage of trimers observed in the experiment, the trimer formation energy should be at least 0.65 eV (we refer to a simulation at 600 K assuming thermodynamic equilibrium). We note that the trimer concentration observed in the experiment at 570 K was $\sim$ 80%.

[1] Long enough to achieve convergence in terms of total energy.

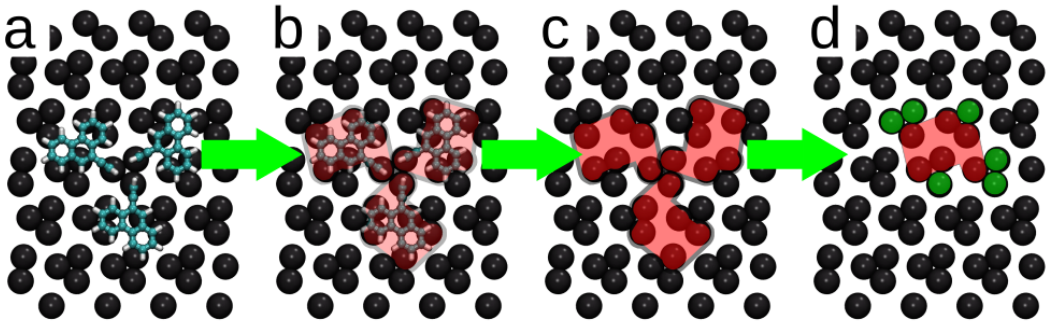


FIGURE 4.8: Obtaining a simplified model of the 9-EP molecule used in MC simulations. (a) Atomic structure of 9-EP molecule adsorbed on the Pd_3 surface. (b) Atomic structure of the 9-EP molecule with a contour overlapping it. (c) Contours of the 9-EP molecule emphasizing the adsorption sites on the Pd_3 surface. (d) Occupied adsorption sites (red) vs free adsorption sites surrounding the molecule (green).

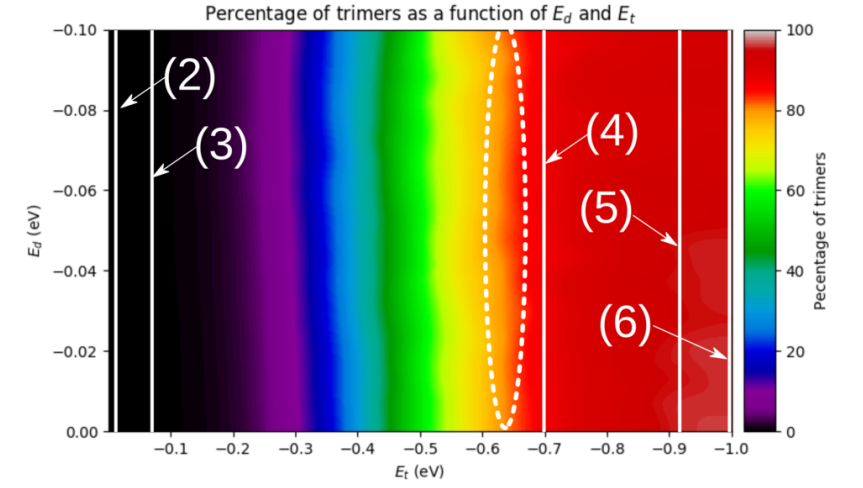


FIGURE 4.9: Percentage of molecules forming trimers as a function of dimer's and trimer's formation energy. The white dotted ellipse is the region that corresponds to experimental observations. The white lines are the DFT-computed formation energies for hypotheses 2-6.

Therefore, it was concluded that the smaller E_t (in absolute value) necessary to explain the trimers' stability is 650 meV. This energy will be further used in the text as a reference to test the remaining proposed hypotheses (2-6).

4.3 Stabilization of the 9-EP trimer by intermolecular forces

The high formation energy derived for the trimers through MC simulations deserves further investigation. In Figure 4.10 we compare the experimental STM image (b) of

the trimer with the simulated one in the case of three intact 9-EP molecules (a) and in the case of removal of H of the ethynyl group (c).

The excellent matching between (a) and (b) with reproduction of the bright fetaure at the center of the trimer, originating from upwards flipping of the three H atoms of the ethynyl groups of the 9-EP molecules, would suggest that the structure of the trimer is the one depicted in (a).

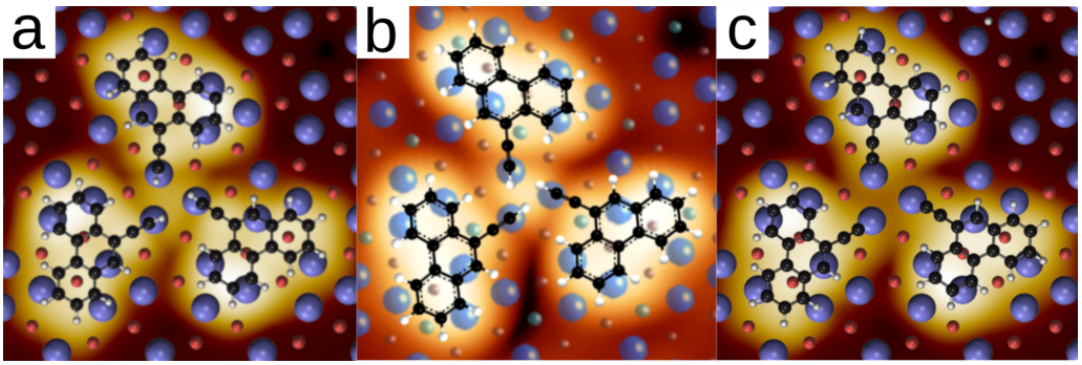


FIGURE 4.10: Simulated images of different hypotheses of the structure of the 9-EP trimer in comparison with the experimental STM. (a) Simulated STM image of the trimer formed by three intact 9-EP molecules. (b) Experimental STM image with an overlap of the 9-EP ball-stick structures. (c) Simulated STM image of the trimer formed by three 9-EP molecules with the dehydrogenated ethynyl group. *Experimental image courtesy of Samuel Stolz, Roland Widmer and Oliver Gröning from the nanotech@surface laboratory at Empa.*

However, DFT calculations of the interaction energy of the three 9-EP molecules when forming a trimer[2] provide a binding energy of just 87 meV per trimer, which is not compatible with the frequency of trimers observed in the experiments. Therefore, hypothesis (2) (see Figure 4.6) was also diregarded.

If we now focus on the trimer of panel (c) where one hydrogen atom per molecule was removed, although the mobility of the trimers would be hindered by the stronger C-Pd binding, the formation energy drops to just 12 meV per trimer. Moreover, it is evident from the simulation that hypothesis (3) (see Figure 4.6) can be excluded due to the absence of the characteristic bright spot at the center.

The results discussed so far draw our attention towards adatoms as a possible explanation of the high formation energy of the trimers. This idea is considered in more detail in the following sections.

[2]Computed as the difference in total energy between three molecules forming a trimer and molecules put far apart.

4.4 Stabilization of the 9-EP trimer by Pd adatom(s)

To study the possible role played by Pd adatoms in stabilizing the trimers, we performed a series of dedicated calculations and experiments. Our experimental collaborators observed that the Pd adatom adsorbed on a Pd_3 surface always binds to the Ga trimers (see Figure 4.11). Moreover, according to the experiment, deposition of Pd atoms on the surface did not affect the occurrence of trimers. This observation suggested that additional Pd atoms may not be the reason for stabilizing the 9-EP trimers on the Pd_3 surface. We now show that DFT simulations support the experiments in excluding the hypothesis of the Pd adatoms as a stabilizing factor for the trimers.

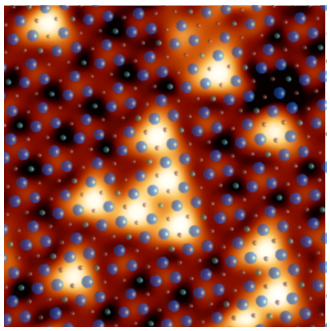


FIGURE 4.11: Experimental image obtained after a deposition of Pd adatoms on the Pd_3 surface. The image is overlapped with a ball model of the first three layers of the Pd_3 surface, where the blue balls represent Pd atoms and the red balls represent Ga atoms. *Experimental image courtesy of Samuel Stolz, Roland Widmer and Oliver Gröning from the nanotech@surface laboratory at Empa.*

DFT geometry optimization, starting from a configuration with a Pd adatom on top of the surface of a Pd trimer, also reveals that the adatom escapes the initial geometry and gets trapped in a local minimum close to a Ga trimer (see Figure 4.12). It

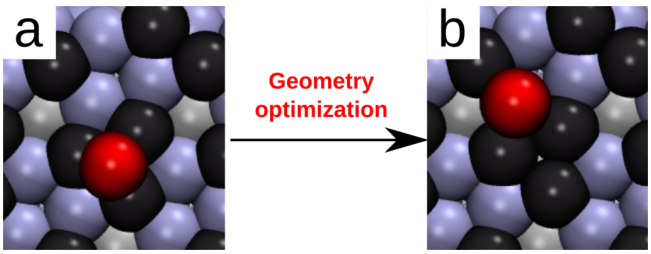


FIGURE 4.12: The Pd atom is not stable in the center of the Pd trimer: upon DFT geometry optimization, it escapes the initial geometry in favor of a local minimum close to the Ga trimer.

is interesting to note that also in the case of aPd adatom initially placed in the center of a 9-EP trimer, the former escapes its starting configuration to reach a neighboring Ga trimer (see Figure 4.13a,b). Final configuration gains 120 meV of energy with respect to the configuration of the Pd atom being in a remote Ga trimer. The total energy gain that could be achieved placing two additional Pd adatoms in symmetry-equivalent positions is 360 meV, which is still far from the targeted 650 meV. Therefore, hypothesis (4) of the Pd-stabilized intact 9-EP trimer was also excluded.

In the case of a trimer formed by 9-EP molecules with a dehydrogenated ethynyls group, the adatom gains $\sim$ 700 meV of energy with respect to Pd adatom placed away from the molecules. Despite this configuration being energetically convenient, the resulting geometry did not reflect the three-fold symmetry observed in the experiments, as the Pd atom got trapped in-between two molecules (see Figure 4.13c,d).

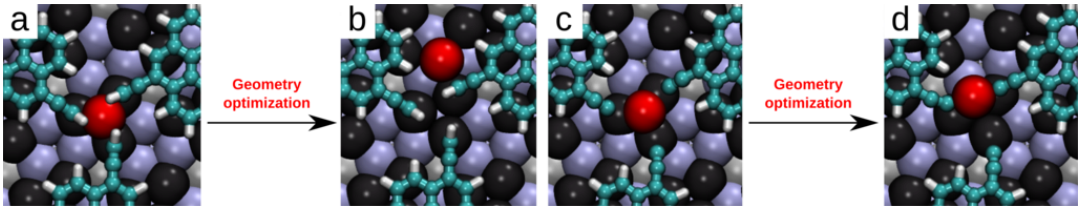


FIGURE 4.13: The Pd atom is not stable in the center of the 9-EP trimer. In the case of a hydrogenated trimer, the Pd atom escapes its initial position in the center (a) and reaches a neighboring Ga trimer (b). In the case of a dehydrogenated trimer (c), the Pd atom is stuck between two 9-EP molecules (d).

Concluding this section, we note that the metal adatom placed in the center of the 9-EP trimer provides significant energetic stabilization. Even though hypothesis (4) considered here (see Figure 4.6) turned out not to be compatible with experimental observations, based on these results we decided to perform the following test to investigate the behavior and STM/nc-AFM appearance of a Ga adatom placed in the center of the 9-EP trimer.

4.5 Stabilization of the 9-EP trimer by a Ga adatom

Experimentally it was observed that Ga adatoms mostly adsorb on two sites of the Pd_3 surface Figure 4.14, one of those being at the center of a Pd trimer. We performed geometry optimizations of Ga adatoms placed in the center of an intact 9-EP

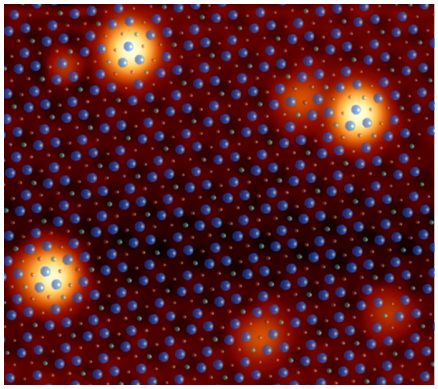


FIGURE 4.14: Experimental STM image of Ga adatoms on the Pd_3 surface. Two different adsorption sites can be distinguished: on top of the Pd trimer and in-between two Pd trimers. *Experimental image courtesy of Samuel Stolz, Roland Widmer and Oliver Gröning from the nanotech@surface laboratory at Empa.*

trimer (see Figure 4.15a) and on a remote Pd trimer. The computed energy difference was $\sim$ 1 eV in favor of the latter, revealing that a Ga adatom is not stable while surrounded by three intact 9-EP molecules. The destabilization due to the repelling interactions between the Ga adatom and the ethynyl groups can be noticed in Figure 4.15a. On the other hand, when considering a 9-EP trimer where ethynyl hydrogens were removed, moving a Ga adatom from an isolated Pd trimer to the center of a 9-EP trimer resulted in an energy gain of 0.9 eV.

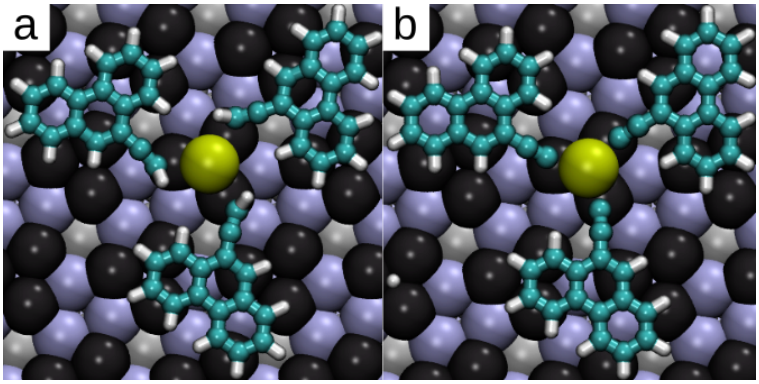


FIGURE 4.15: Optimized 9-EP geometries with a Ga adatom in the center of hydrogenated (a) and dehydorgenated (b) trimers. It can be noticed that in the first case the ethynyl groups tend to avoid the Ga adatom, while in the second case the dehydrogenated ethynyl groups bind to it.

The above energetic considerations, together with a good agreement of simulated and experimental STM images (see Figure 4.16), lead to a hypothesiss (5) of dehydrogenated 9-EP trimers with a Ga adatom at the center compatible with the

experimental STM images for the 9-EP trimers.

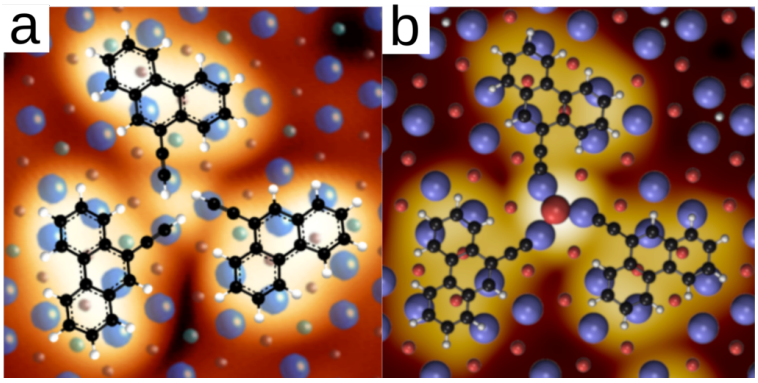


FIGURE 4.16: Comparison of an experimental image of 9-EP trimer with a simulated STM image of dehydrogenated 9-EP molecules with an Ga adatom . (a) Experimental STM image. (b) Simulated STM image of a dehydrogenated 9-EP trimer with Ga adatoms in its center. *Experimental image courtesy of Samuel Stolz, Roland Widmer and Oliver Gröning from the nanotech@surface laboratory at Empa.*

To test it further we performed an additional check: we compared the total energy difference between 3 intact 9-EP molecules forming the trimer with a Ga adatom placed on an isolated Pd trimer and 3 dehydrogenated 9-EP molecules connected to the Ga adatom. The energy of the latter is higher by $\sim$ 1.04 eV, which is very significant and, in principle, does not allow to accept hypothesis (5). But rigorous thermodynamic considerations should also include the entropic factor, which acts in favor of losing a hydrogen atom: once the atom is detached, it is quite unlikely that it can find a way to attach back. Moreover, as it was discussed in section 4.1, experimental observations suggest that the 9-EP molecule indeed starts to lose hydrogen atoms already at 400 K.

Therefore, we conclude this section by highlighting the fact that among all hypotheses analyzed so far, stabilization of the 9-EP trimer by a Ga adatom matches very well both the experimental and theoretical results.

4.6 Stabilization of the trimer by CO molecule

Our final investigation considers CO molecules that, as observed by Prinz et al. [14], adsorb in the center of Pd trimers on the Pd_3 surface. In Figure 4.17 we show different geometries that were tested for CO molecules interacting with 9-EP trimers. CO molecules in the center of intact trimers (see Figure 4.17b) increase the energy by

0.9 eV with respect to CO on isolated Pd trimers. On the other hand, formation of a tertiary alcohol (Figure 4.17a) lowers the energy by $\sim$ 1 eV. In analogy to the case of Ga adatoms, we also considered CO molecules at the center of dehydrogenated 9-EP trimers (Figure 4.17c) . The energy of the latter is higher by $\sim$ 1.2 eV compared to

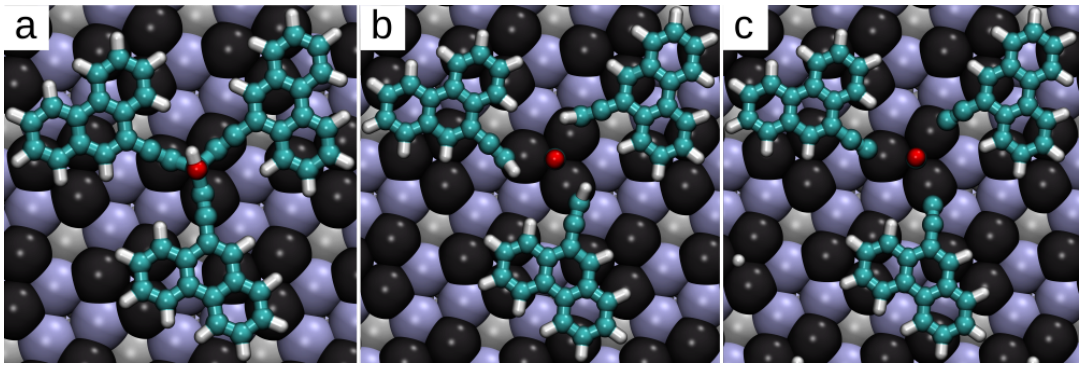


FIGURE 4.17: Optimized geometries obtained for a 9-EP trimer with a CO molecule in the center. (a) Tertiary alcohol in the center of a 9-EP trimer. (b) CO molecule in the center of an intact 9-EP trimer. (c) CO molecule put in the center of a dehydrogenated 9-EP trimer.

.

a dehydrogenated trimer and a CO molecule on a free Pd trimer far from 9-EP. This is due to the fact that the CO being in the center interacts weakly with the surface Pd atoms already saturated by 9-EP radicals; this is also evident from the increased height of CO, which is 1.8 Å higher than on the free Pd trimer. We also provide simulated STM images for comparison with the case of tertiary alcohol (see Figure 4.18a) and experiment (see Figure 4.18b).

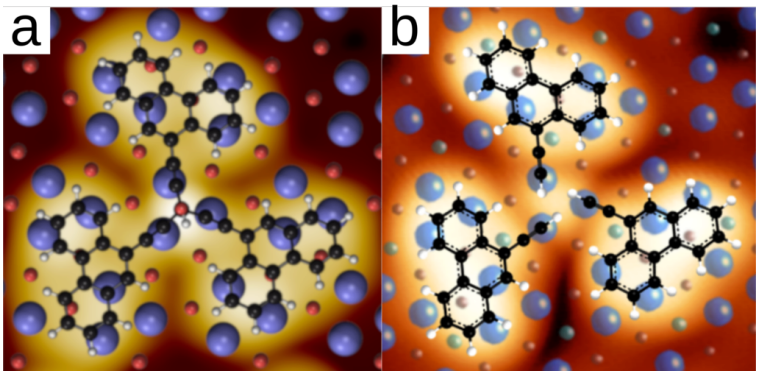


FIGURE 4.18: Comparison of a simulated STM image of a 9-EP trimer with tertiary alcohol formed in the middle (a) with an experimental STM image (b). *Experimental image courtesy of Samuel Stolz, Roland Widmer and Oliver Gröning from the nanotech@surface laboratory at Empa.*

Formation of covalent bonds with the central carbon atom of the CO molecule

implies that all the ethynyl groups are oriented towards the same central atom (see Figure 4.18a). This is not compatible with the experimental observations where the ethynyl groups are oriented to neighboring Pd atoms in the Pd trimer (Figure 4.18b). Since STM images did not allow to unambiguously assess the orientation of the ethynyl groups, to further investigate this configuration our experimental collaborators performed nc-AFM experiments allowing to unravel the geometry of 9-EP molecules. In Figure 4.19 we show a comparison between the experimental and theoretical nc-AFM images obtained for different molecular models.

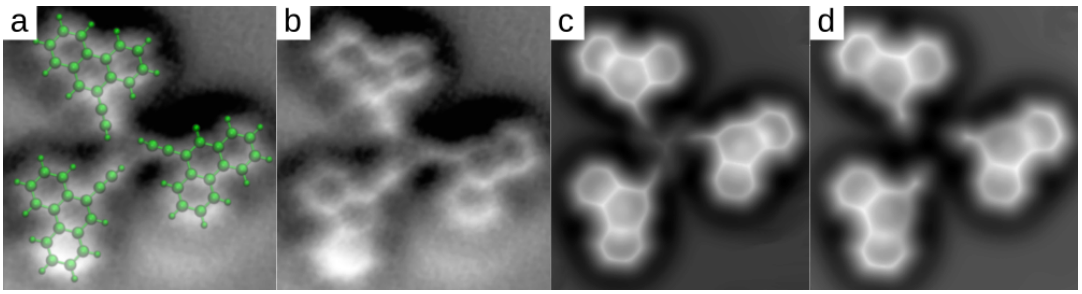


FIGURE 4.19: Comparison of experimental and simulated nc-AFM images for 9-EP trimers. (a) Experimental AFM image and superimposed molecular sketches. (b) Same experimental nc-AFM image. (c) Simulated nc-AFM image of the 9-EP with Ga adatom in the center of the trimer. (d) Simulated nc-AFM image of the trimer formed by intact 9-EP molecules. *Experimental image courtesy of Samuel Stolz, Roland Widmer and Oliver Gröning from the nanotech@surface laboratory at Empa.*

According to the experiment, ethynyl groups are indeed not oriented to the center of the trimer as can be seen from the overlap of experimental nc-AFM images with atomic sketches (Figure 4.19a). This is in agreement with the simulated images for the case of Ga adatoms in the center of a dehydrogenated trimer, even though no features at the center of the trimer are evident in the experiment (Figure 4.19c). This latter discrepancy could depend on poor parameterization of the force field used in nc-AFM simulations for the case of a Ga atom. The ethynyl groups of a 9-EP intact trimer look much shorter (Figure 4.19d) in comparison with the experimental AFM image, which again suggests that hypothesis (2) is not valid. We do not include the simulation images of nc-AFM for the case of tertiary alcohol since they end up being considerably distorted due to the fact that our simulation approach for nc-AFM does not allow for substrate relaxation.

Concluding this section, we can say that the CO molecule in the center can not stabilize the 9-EP trimer in either of the discussed ways: being in the center of the intact/dehydrogenated trimer (because of the significant energy increase) or forming covalent bonds (as it results from the nc-AFM analysis). Therefore, we conclude that only hypothesis (5) can explain the stability of the 9-EP trimer because it matches all the geometrical, energetic and spectroscopic criteria.

4.7 R ↔ S transformation mechanism

We conclude the study of the Pd_3 surface investigating the "flipping" mechanism (to identify the mechanism of $R \leftrightarrow S$ transformation) of 9-EP molecules at the surface in analogy with what was done for the Pd_1 case. For the Pd_3 surface we obtain a barrier of 1.2 eV (see Figure 4.20), which is slightly higher in comparison with the Pd_1 case (1.1 eV) but still compatible with the experimental observations that the transformation happens already at 400 K.

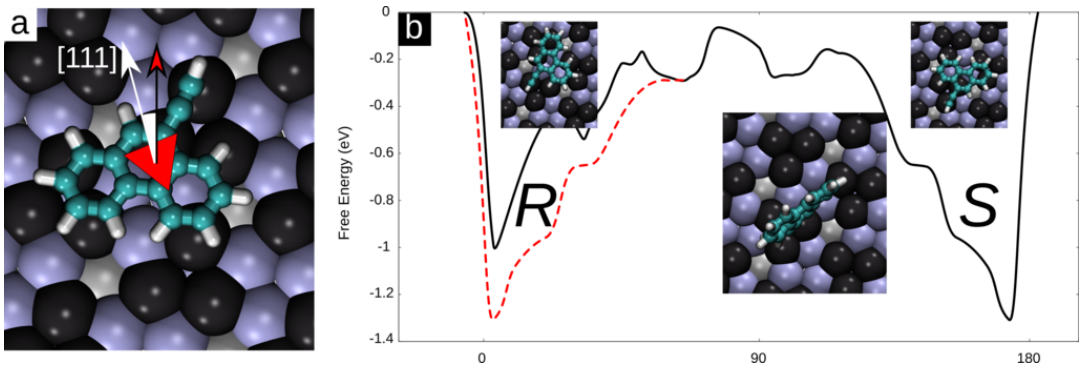


FIGURE 4.20: Flipping of a 9-EP molecule adsorbed on the PdGa:A($\bar{1}\bar{1}\bar{1}$)Pd_3 surface. (a) Description of the collective variable used in the metadynamics (b) Approximate free energy landscape obtained after the first crossing has occurred; the barrier for the flipping is 1.2 eV. The red dashed line is a guide to the eye mimicking the second basin corresponding to the R configuration. Superimposed are atomistic sketches of the most relevant configurations for the 9-EP molecule along the flipping path.

Even though it has been discussed in this section that a 9-EP molecule on the Pd_3 surface may be dehydrogenated, we restricted our consideration to an intact molecule only. Since during the flipping the molecule keeps its ethynyl group attached to the Pd_3 surface, we are confident that the hydrogenation state of the molecule does not have a major effect on the "flipping" energy.

4.8 Methods

All the DFT simulations discussed here were performed by means of the CP2K code [96, 97] implementing DFT within a mixed Gaussian plane waves approach [98]. The surface/adsorbate systems were modeled within the periodically repeated slab scheme [155] and the Martyna-Tuckerman approach [87] was employed to decouple the surfaces from their replica. The electronic states were expanded with the TZV2P Gaussian basis set [156] for C and H atoms and the DZVP basis set for Pd and Ga. A cutoff of 1200 Ry was used for the plane waves basis set. Goedecker-Teter-Hutter pseudopotentials [157] were used to represent the frozen core of the atoms. Eighteen and thirteen electrons were included in the valence for Pd and Ga species respectively.

For all the simulations (except methadynamics) the model system contained 23 atomic layers of PdGa along the [$\bar{1}\bar{1}\bar{1}$] direction, the adsorbate and 40 Å of vacuum. We used the revPBE [148] exchange correlation functional and DFT-D3 model [69] for van der Waals interactions. To optimize the adsorption geometries, we considered supercells of 20.76 $\times$ 23.96 Å (corresponding to twelve surface units) to study monomers and 27.67 $\times$ 35.95 Å (corresponding to twenty four surface units) to study trimers. In both cases, the size of the unit cell in Z direction was 60 Å. We kept the atomic positions of the bottom 6 layers of the slab fixed to the ideal bulk positions, all other atoms were relaxed till forces were lower than 0.005 eV/Å.

To obtain simulated STM images [126, 127], within the Tersoff-Hamann approximation [158, 125], we extrapolated the electronic orbitals to the vacuum region following the approach outlined by Tersoff [125].

For metadynamics, we considered 7 atomic layers (approx. 5 Å) along the [$\bar{1}\bar{1}\bar{1}$] direction, the adsorbate and 20 Å of vacuum. The supercell size was 13.83x23.96 Å, corresponding to eight surface units. We kept the atomic positions of the bottom 3 layers of the slab fixed to the ideal bulk positions. This allowed us to perform trajectories of approximately 25 ps to study the flipping process.

For the nc-AFM simulations we used the ProbeParticle model [33] implemented in the two-point version (see chapter 6) to account for the Carbon and Oxygen atom of the CO molecule.

4.9 Conclusion

Studying the adsorption of 9-ethynylphenanthrene on the Pd_3 surface turned out to be a challenge. Having identified the geometry of the 9-EP monomers on the surface, we could not unambiguously determine its hydrogenation state. Even though it would be possible to say that at room temperature the molecules stay intact, it is quite likely that at temperatures higher than 400 K the 9-EP molecules lose hydrogen atoms of the ethynyl group.

Another challenge was to determine the origin of the stability of 9-EP trimers. Since their formation is a process which goes against thermodynamically favorable increase of entropy, it could only be explained by an interaction energy high enough to compensate for the entropy loss. We tested several hypotheses that could explain the origin of such stabilization, and the "closest" hypothesis we could identify was stabilization of a 9-EP trimer by a Ga adatom in the center of it. The simulated STM, as well as the simulated nc-AFM images of this model, were in good agreement with the experiment. However, with consideration to energy, it could only be accepted if the ethynyl group of the 9-EP molecules were initially dehydrogenated, which seemed to happen at 400 K.

Other tested hypotheses could be dismissed. The formation of a benzene ring was found to be sterically impossible. The intermolecular energy was found to be lower than the entropy contribution so could not stabilize the trimer either. The only hypothesis that remained to be clarified was the stabilization of trimers by chemical transformation. The natural candidates to be placed in the center were Pd and Ga adatoms. Experimentally and computationally, it was shown that a Pd adatom is not stable in a Pd trimer, so we discarded this hypothesis. The 3 Pd atoms placed on top of the Ga trimers that surround the central Pd trimer did not provide enough stabilization energy. The CO molecule placed in the center could not stabilize neither the dehydrogenated nor the intact trimer. Formation of tertiary alcohol could be an option as well; however, according to the experimental nc-AFM image the ethynyl groups do not point to the same central, which would obviously be the case if alcohol had been formed.

According to the DFT-based metadynamics simulations, the $R \leftrightarrow S$ transformation mechanism is the same as for the Pd_1 surface: flipping of the 9-EP molecule with an estimated free energy barrier not higher than 1.2 eV.

Chapter 5

Adsorption of Hexabenzocoronene on Pt(111) surface

Following the introduction given in the section 1.4, the study of Hexabenzocoronene (HBC) adsorbed on a Pt(111) surface will be discussed here in more detail. This work was done combining scanning tunneling microscopy (STM) experiments with density functional theory (DFT) simulations. The preferred adsorption motifs for HBC when varying the coverage were investigated. It was found that at low exposure, an isolated HBC has two symmetry-equivalent adsorption positions: the molecule is located on the top site and the angle between the long-symmetry edge of HBC with the Pt(111) surface lattice directions is $\pm(10 \pm 2)^{\circ}$. This result was confirmed by DFT simulations, where the theoretically determined angle was $\pm 11.0^{\circ}$. Furthermore, these two structures have shown a significant energetic favor (in the order of 1 eV) with respect to the other tested positions: fcc-hollow, hcp-hollow and bridge.

Further, the formation of HBC domains at higher coverage was explored. Using STM we showed that HBC forms two types of domains adsorbed on Pt(111) with an intermolecular distance of 16.3 and 17.3 Å. Interestingly, as it was demonstrated in the literature [162], HBC adsorbed on Cu(111) and Au(111) creates more compact domains with a nearest neighbor distance of 14.2 Å and 15.0 Å respectively. By means of the DFT simulations, we reveal that in the case of HBC adsorbed on the Pt(111) surface the Pauli repulsion plays the dominant important role, while in the case of HBC adsorbed on Cu(111) and Au(111) the intermolecular attraction was

mostly attributed to vdW interactions. For this reason, HBC adsorbed on Cu(111) and Au(111) metals keeps the smallest possible distance between molecule units, while on Pt(111) the molecules try to keep themselves further apart.

5.1 Low HBC coverage on a Pt(111) surface

We started our investigation from the adsorption geometry of HBC on Pt(111) at low coverage. The experimental STM image in Figure 5.1 shows separated HBC molecules, indicating a repulsive interaction between them. Our experimental collaborators have identified two distinct adsorption geometries **A** and **B** that are presented in Figure 5.1b,c.

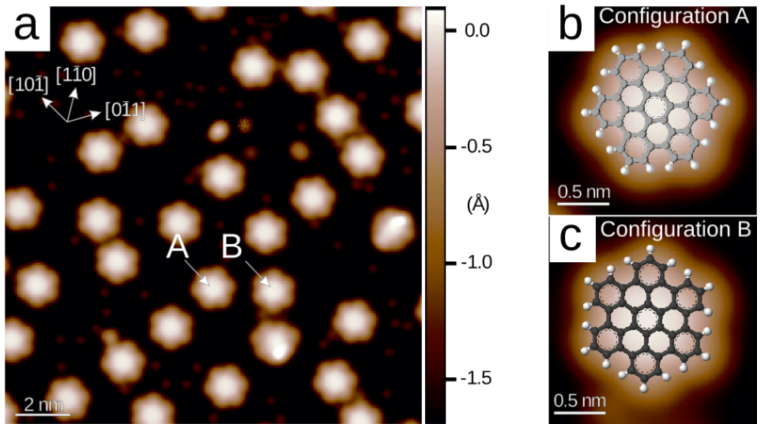


FIGURE 5.1: (a) STM image (10 nm × 10 nm) of HBC adsorbed on Pt(111) surface. High symmetry crystal directions are indicated. A closer view (2 nm × 2 nm) of the areas indicated by the arrows in (a). (b) Shows a closer view of the **A** configuration. (c) Shows a closer view of the **B** configuration. For better representation, the STM image in (a) and (b) is superimposed with ball-stick models of the HBC molecule. *Experimental image courtesy of Yves Forrer and Oliver Gröning from the nanotech@surface laboratory at Empa.*

To understand the details of the equilibrium geometries we performed DFT calculations of the HBC molecule centered on four possible inequivalent sites on the Pt(111) surface: top (**T**), hcp-hollow ($\mathbf{H_h}$), fcc-hollow ($\mathbf{H_f}$) and bridge (**Bg**) that are marked in Figure 5.2c. For each adsorption site, considering the six-fold symmetry of HBC, we tested different initial orientations of the molecule in the range $[-20°; 30°]$[1]. The angle was measured between the $[1\bar{1}0]$ crystallographic axis and vector connecting the center of the HBC molecule with hydrogen atom at the corner of the

[1]Please note that -30° and 30° would be equivalent due to the six-fold symmetry of the HBC molecule.

molecule (see Figure 5.2a — green and red arrows and Figure 5.2b — green and blue arrows). Optimizing the 24 geometries (4 sites $\times$ 6 rotations from -30° to 20°) we found out that the two most stable ones are located on the **T** site. Their structures are shown in Figure 5.2a,b.

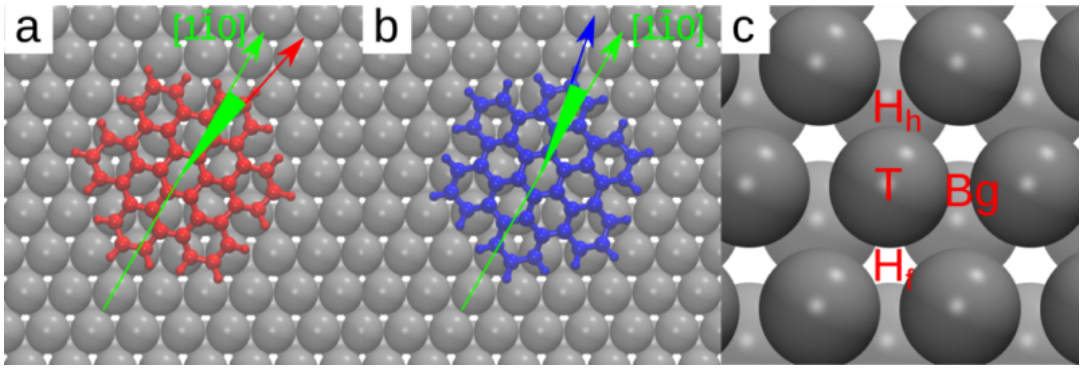


FIGURE 5.2: The two most stable (according to DFT) orientations of a HBC molecule adsorbed on a Pt(111) surface on the **T** adsorption site. Orientation **A** is shown in panel (a), orientation **B** is shown in panel (b). The gray spheres represent Pt atoms. Carbon and hydrogen atoms are drawn using the same color: red for the **A** enantiomer and blue for the **B**. Carbon and hydrogen atoms can be distinguished based on the size (hydrogen is drawn smaller). The mirror plane distinguishing **A** and **B** configurations is perpendicular to the (111) surface and goes along the green arrow representing the [1$\bar{1}$0] crystallographic axis. (c) Representation of different adsorption sites investigated in the current study: **T**- top site, $\mathbf{H_h}$ - hcp-hollow site, $\mathbf{H_f}$ - fcc-hollow site, **Bg** - bridge site.

The energies obtained after optimization are summarized in Table 5.1. Four minima are present with adsorption energy of -8.00 eV for **T** site starting from orientations -20°, -30°, 10°, 20°. These four minima correspond to two distinct adsorption geometries: upon relaxation -10°, -20° lead to the same endpoint with $\alpha_{DFT} = -11.0^\circ$ (Figure 5.2a) and 10° , 20° lead to $\beta_{DFT} = 11.0^\circ$ (Figure 5.2b).

The geometry found in experiments and labeled as **A** is characterized by an angle $\alpha_{exp} = -10^\circ \pm 2^\circ$ between the [0$\bar{1}$1]-direction (indicated in Figure 5.2a) and the nearest corner of HBC. The same angle of geometry **B** is $\beta_{exp} = 9^\circ \pm 2^\circ$, where the uncertainty arises from the identification of the midpoint of the molecule and from the assignment of the [0$\bar{1}$1] axis. As it can be seen, the corresponding angles obtained from the experiments and from DFT match each other quite well, therefore we concluded that the identified atomic models of HBC adsorbed on the Pt(111) surface agree with the experimental observations.

TABLE 5.1: Adsorption Energies (in eV) for HBC on Pt(111) for the optimized geometries obtained starting from 24 different guesses. The final angle describing the molecule's orientation is shown in the parentheses. The configurations which led to the most stable domains are highlighted in yellow.

Starting angle	**T**	$\mathbf{H_h}$	$\mathbf{H_f}$	**Bg**
30.0°	−7.21 (30.2°)	−6.98 (30.0°)	−6.92 (30.1°)	−7.00 (30.2°)
20.0°	−7.99 (11.1°)	−6.97 (29.8°)	−6.87 (29.8°)	−7.03 (30.0°)
10.0°	−7.99 (11.0°)	−6.89 (−0.4°)	−6.52 (8.6°)	−6.93 (3.9°)
0.0°	−6.43 (0.0°)	−6.90 (−0.4°)	−6.38 (0.0°)	−6.91 (0.4°)
−10.0°	−7.99 (−11.0°)	−6.90 (−0.4°)	−7.99 (−11.7°)[a]	−6.93 (0.3°)
−20.0°	−7.99 (−11.0°)	−7.15 (−29.7°)	−6.87 (30.0°)	−7.03 (−29.1°)

[a] The structure initially placed on the $\mathbf{H_f}$ position along the DFT optimization shifted to the **T**.

5.2 High HBC coverage on a Pt(111) surface

The next step from the experimental side was to analyze adsorption motifs in the high coverage regime. HBC was deposited up to 70% of a complete monolayer and characterized the molecular domains by means of STM. In Figure 5.3 an experimental STM image of the HBC monolayer on Pt(111) is shown together with insets that allow to identify peculiarities of different domains. The inset presents chemical sketches of HBC superimposed on the STM images (Figure 5.3b-e).

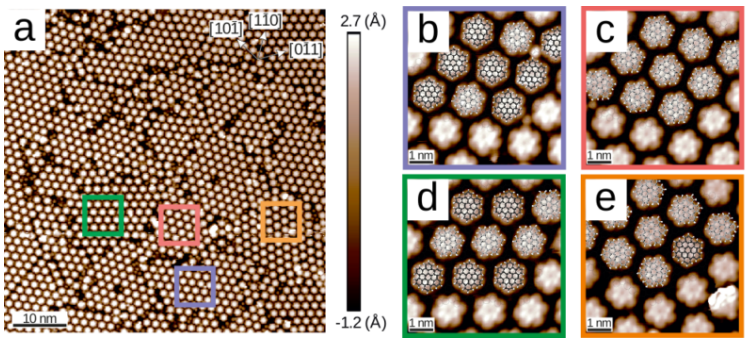
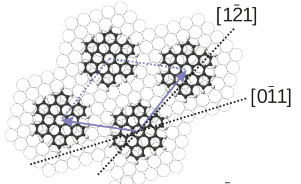
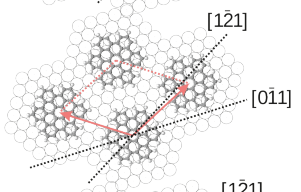
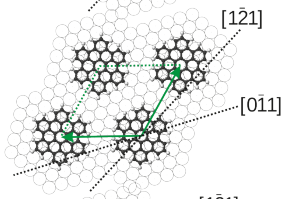
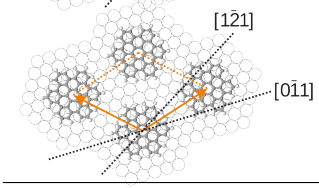


FIGURE 5.3: (a) STM image (60 nm × 60 nm) of a 70% HBC ML covered Pt(111) surface. (b–e) Closer views (6.6 nm × 6.6 nm) of the colored insets marked in (a). *Experimental images courtesy of Yves Forrer and Oliver Gröning from the nanotech@surface laboratory at Empa.*

Combining the information obtained for the adsorption geometries at low coverage together with the microscopy information obtained at high coverage, the repeat units proper to each domain can be identified as listed in Table 5.2 where we define lattice parameters for each one of the four domains highlighted in Figure 5.3.

TABLE 5.2: Structural data for the HBC domains on Pt(111) observed at a HBC coverage of 70% ML. Please notethe colors of the arrows representing the unit cells. These colors are preserved in the whole chapter to identify the corresponding domains. *Analysis of the experimental results was performed by Yves Forrer under the supervision of Oliver Gröning from the nanotech@surface laboratory at Empa.*

Domain	Reference	Adsorption geometry	Lattice parameters	Epitaxy matrix
	(7, 4)	preferential **B**	$a = \sqrt{37} * a_{Pt(111)}$ $\gamma = 120°$	$\begin{bmatrix} 7 & -4 \\ 4 & 3 \end{bmatrix}$
	(7, 3)	preferential **A**	$a = \sqrt{37} * a_{Pt(111)}$ $\gamma = 120°$	$\begin{bmatrix} 7 & -3 \\ 3 & 4 \end{bmatrix}$
	(7, 5)	preferential **B**	$a = \sqrt{39} * a_{Pt(111)}$ $\gamma = 120°$	$\begin{bmatrix} 7 & -5 \\ 5 & 2 \end{bmatrix}$
	(7, 2)	preferential **A**	$a = \sqrt{39} * a_{Pt(111)}$ $\gamma = 120°$	$\begin{bmatrix} 7 & -2 \\ 2 & 5 \end{bmatrix}$

To identify the surface areas where each domain occurs, we applied Fast Fourier Transform (FFT) to the STM image (Figure 5.3a), thus mimicking the result of a LEED[2] experiment. Knowing the structural information about all the most important surface patterns (as indicated in Table 5.2), we could assign them to each point in Figure 5.3a, which gives the picture shown in Figure 5.3b. A backwards FFT of the latter LEED pattern for each specific color allows to retrieve the surface area occupied by the corresponding domain as shown in Figure 5.4c, where the back-Fourier transformed image is superimposed with the original STM one.

[2]Low energy electron diffraction.

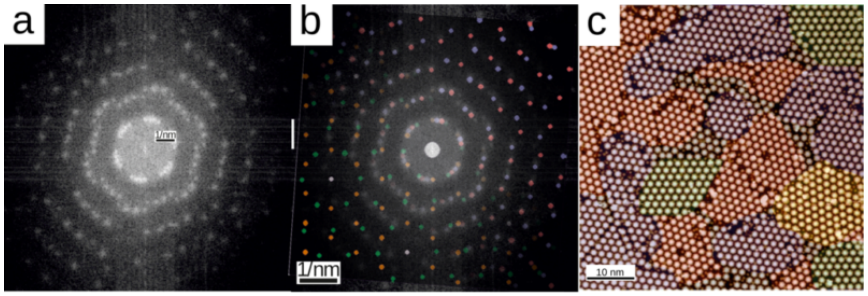


FIGURE 5.4: (a) Simulated LEED image obtained from a 2D-FFT of Figure 5.3a. (b) Assignment of LEED spots to the domains shown in Figure 5.3b-e and Table 5.2. (c) Domain assignment obtained by means of inverse FFT of (b) the colors are consistent with the colors of the arrows defining the domains in the Table 5.2. *Simulated image courtesy of Yves Forrer and Oliver Gröning from the nanotech@surface laboratory at Empa.*

The image clearly shows that at high coverage the Pt(111) surface is almost exclusively occupied by four kinds of domains: (7, 2), (7, 3), (7, 4) and (7, 5) while (6, ...) domains were not formed on the surface. To understand this we again employed DFT calculations to rank the stability of the surface domains. This will be discussed in more details in the next section.

5.3 Origin of the intermolecular repulsion

To start the discussion about the instability of the close packed (6, ...) domains and the stability of (7, ...) ones we first summarize their structural information in one picture. In the Figure 5.5, domains are represented highlighting the occupied sites of the Pt(111) surface. To construct a particular domain, one should put the first molecule in the center of the picture (highlighted in black) and surround it by another six molecules to the colored spots corresponding to the domain of interest.

It is also useful to discuss the symmetry of the domains. Looking carefully at (7, 2) formed by the molecule in **B** configuration and at (7, 5) formed by the molecule in **A** configuration, one could notice that they are identical (see Table 5.2). The same is true for (7, 3)-**A** and (7, 4)-**B**, for (6, 2)-**A** and (6, 4)-**B**. Naturally, if one replaces **A** with **B** and *vice versa* – those statements would still remain valid. Therefore, it is enough to investigate the domains having the molecules in one configuration since due to the symmetry they will also cover another configuration. We had arbitrarily

chosen to study **B**, so further in the text only a **B** configuration on the **T** site will be assumed unless otherwise specified.

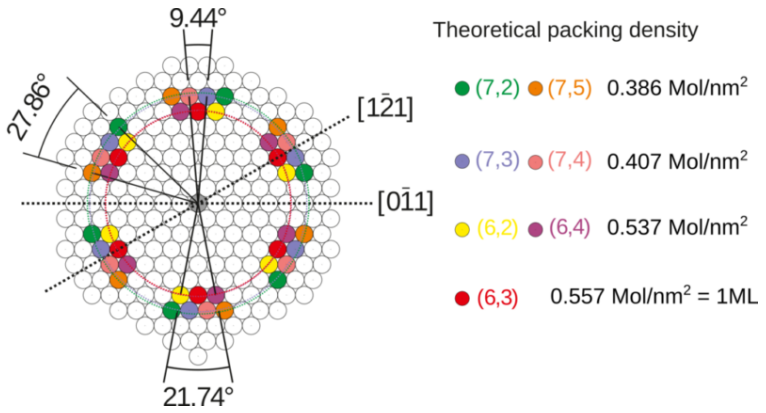


FIGURE 5.5: Representation of the domain structures with their reference (x,y) and corresponding color. The packing density is also provided. To construct a domain, one should place an HBC molecule on the central black spot and surround it by another six molecules placed to the spots highlighted with the corresponding domain color. *Image courtesy of Yves Forrer and Oliver Gröning from the nanotech@surface laboratory at Empa.*

As the starting point we considered HBC pairs in vacuum having planar monomer's geometry and forming all the domains indicated in the Figure 5.5. The results, listed in Table 5.3, show (6, 3) as the most stable one in clear contradiction with the experimental evidence where no (6, ...) domains were reported. Since the HBC-HBC interaction in a vacuum is dominated by vdW contributions, it was concluded that they do not play the key role in the stabilization.

TABLE 5.3: Relative energies of HBC dimers in vacuum placed in positions corresponding to the experimentally observed configurations. All energies are in meV with "0" corresponding to the most stable dimer.

Domain	Relative energy (meV)	Distance between centers (Å)
(6, 2)	19.8	14.68
(6, 3)	0.0	14.42
(6, 4)	44.4	14.68
(7, 2)	96.6	17.33
(7, 3)	94.0	16.88
(7, 4)	92.7	16.88
(7, 5)	93.7	17.33

As the next step to understand the experimental evidence of the stability of the (7, ...) domains, we added Pt substrate assuming that its presence may have an influence. We again computed the DFT energies of the same domains listed in the Table 5.3 but now they were optimized on the Pt(111) surface. In Table 5.4 we report the energy values with respect to the most stable (7, 2) domain. From the DFT simulations it is clear that the total energy of (7, 2) and (7, 3) dimers is lower compared to (6, 2) and (6, 3) ones, which agrees very well with the experimental observations. However, the origin of the instability of (6, ...) domains still has to be clarified. The dispersion interaction played an important role for the domains formed on Cu(111) and Au(111) surfaces [162] also in that case favors the close packed (6, 2) and (6, 3) ones. Same is true for the electrostatic contributions to the total energy (also shown in the Table 5.4), which does not to fit to the experimental evidence either.

The only contribution that goes along the total energy, and, eventually, explains the stability of (7, 2) dimer is the core Hamiltonian energy or kinetic energy of the electrons. The increase of the later is at the origin of the Pauli repulsion, as according to quantum mechanics when two electronic states of different systems start to overlap they have to be orthogonal to each other[3].

TABLE 5.4: Relative total energy and vdW energy values with respect to the most stable (7, 2) domain. All energies are in meV. All the dimers are formed placing the molecules in **B** configuration on **T** adsorption sites.

Position	Total energy	Dispersion energy	Electrostatic energy	Core Hamiltonian energy
(6, 2)	223.1	-68.1	-2190.0	3417.3
(6, 3)	259.0	-75.2	-3091.5	4559.1
(6, 4)	229.6	-60.5	-1518.4	2457.8
(7, 2)	0.0	0.0	0.0	0.0
(7, 3)	25.4	-3.4	-996.2	1369.7
(7, 4)	46.3	-13.8	-427.7	742.1
(7, 5)	38.6	-14.8	-508.8	806.3

Those results reveal a crucial role that Pt substrate plays in the structure of HBC domains. To understand how the substrates influences on the electronic density

[3]While the orthogonalization itself costs energy.

distribution we computed the charge-density difference of a monomer to have a visual representation of the changes that the system undergoes when the molecule adsorbs on the Pt(111) surface

$$\rho(\vec{r}) = \rho_{sub-mol}(\vec{r}) - \rho_{sub}(\vec{r}) - \rho_{mol}(\vec{r}) \tag{5.1}$$

where $\rho_{sub-mol}(\vec{r})$ is the charge-density of the molecule/substrate complex, $\rho_{sub}(\vec{r})$ is the charge-density of the isolated substrate and $\rho_{mol}(\vec{r})$ is the charge-density of the isolated HBC molecule. To compute the charge-density difference, the geometries of the isolated systems were constrained to their geometries in the molecule-substrate complex.

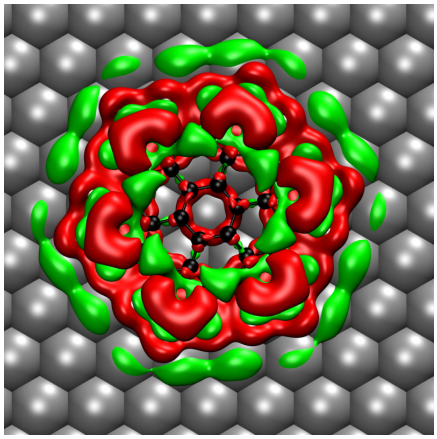


FIGURE 5.6: HBC charge density difference between the molecule/substrate complex and isolated molecule, isolated substrate. The green "clouds" represent an excess of charge-density difference, while the red clouds represent a lack of it. This image indicates a significant decrease of the molecule's electron density and increase of the electron density at the Pt(111) surface.

The visual representation of the charge-density difference suggests that the electron density moves from the molecule to the surface of the substrate. To quantify this observation we plot on the same graph the electron density difference from the side view and the integral of the charge-density (dark plot) within every plane along Z direction, see Figure 5.7. The thickness of every plane is $\sim$ 0.1 Å. The plot also reveals an increase of electron density close to the Pt surface (green area in the plot), and a decrease of the density of the molecule (red area in the plot). In total we obtained a decrease of $\sim$ 1.1 electrons in the red area and an increase of $\sim$ 1.2 electrons

in the green area. The closeness of those two number supports the idea that electron density excess on the Pt surface almost exclusively originates from the HBC molecule.

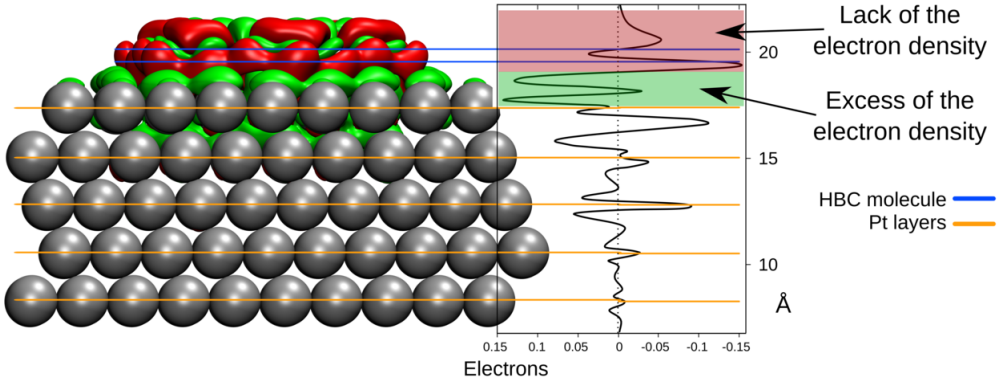


FIGURE 5.7: Plane average of the charge-density (colored in blue) and the integral of the charge-density (colored in red). The distance along z direction is given in Å, the integrals of the charge density are in electrons.

To demonstrate how the transferred electron densities overlap in the case of dimers we computed the charge-density difference between HBC dimer/substrate and isolated HBC dimer, substrate, see Figure 5.8a,b. In this particular case we considered only the most stable (6, 2) and (7, 2) domains, as for the other (6, ...) and (7, ...) the same consideration would remain valid. As it can bee seen in the case of (7, 2) domain the transferred electronic density of two neighboring molecules does NOT overlap, while in the case of the (6, 2) domain it eventually DOES overlap. Unfortunately the Figure 5.8a does really show the changes that are happening in the intermediate region between two molecules, when they approach each other. To demonstrate this we computed the so-called double density difference defined as follows

$$\rho(\vec{r})^d = [\rho_{sub-mol1-mol2}(\vec{r}) - \rho_{sub}(\vec{r})] - [\rho_{sub-mol1}(\vec{r}) - \rho_{sub}(\vec{r})] - [\rho_{sub-mol2}(\vec{r}) - \rho_{sub}(\vec{r})] \quad (5.2)$$

where $\rho_{sub-mol1-mol2}(\vec{r})$ is the charge-density of the HBC-dimer/substrate complex, $\rho_{sub-mol1}(\vec{r})$ is the charge-density of the HBC-monomer1/substrate complex, $\rho_{sub-mol2}(\vec{r})$ is the charge-density of the HBC-monomer2/substrate complex, $\rho_{sub}(\vec{r})$

is the charge-density of the isolated substrate. Here the double density difference $\rho(\vec{r})^d$ shows the changes that two HBC molecules undergo when they approach each other on the surface from infinitely long distance. As it can be seen in the Figure 5.8c in the case of (6, 2) domain there is a clear change of charge-density in the region between the molecules, while in the case of the (7, 2) domain (Figure 5.8d) the changes are significantly less pronounced.

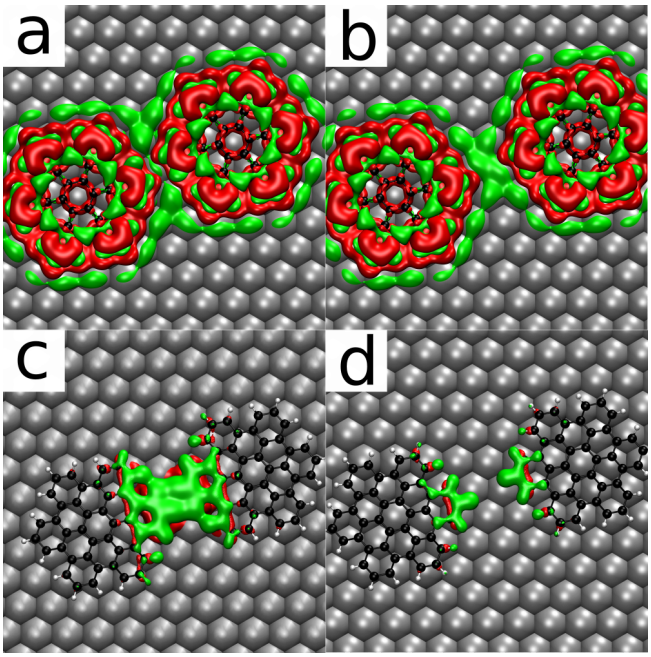


FIGURE 5.8: (a) Charge-density difference $\rho(\vec{r})$computed for the (6, 2) domain. (b) Charge density difference $\rho(\vec{r})$computed for the (7, 2) domain. (c) Double charge-density difference $\rho(\vec{r})^d$computed for the (6, 2) domain. (d) Double charge-density difference $\rho(\vec{r})^d$computed for the (7, 2) domain.

As molecules approach each other, the transferred charge-density (described in the Figure 5.7) of two separate molecules shifts upwards to avoid overlapping. This transformation result in increase of the Pauli energy, making the close-packed (6, ...) domains significantly less stable compared to the (7, ...) ones and thus providing the origin of the stabilization of the (7, 2) domain.

5.4 Methods

All the simulations presented in the chapter were performed by means of the CP2K code [96, 97] implementing DFT within a mixed Gaussian and plane waves approach [98]. The surface/adsorbate systems were modeled within the periodically

repeated slab scheme [155] and the Martyna-Tuckerman approach [87] was employed to decouple the surfaces from their replica. The electronic states were expanded with the TZV2P Gaussian basis set [156] for C and H atoms and the DZVP basis set for Pt. A cutoff of 1200 Ry was used for the plane waves basis set. Goedecker-Teter-Hutter pseudopotentials [157] were used to represent the frozen core of the atoms. Eighteen electrons were included in the valence for Pt species.

For all the simulations, the model system contained 5 atomic layers of Pt along the [111] direction, bottom layer of hydrogen atoms, the adsorbate and 30 Å of vacuum. The PBE exchange correlation functional and Grimme's approach [69] to obtain van der Waals interactions was applied in all the calculations. The bulk lattice parameter for Pt was obtained optimizing the supercell consisting of $4 \times 4 \times 4$ unit cells. To optimize the adsorption geometries of single molecules supercells measuring 27.75x24.03 Å corresponding to 100 surface units were considered. For the dimers supecells measuring 44.39×38.45 Å corresponding to 256 surface units were employed. The atomic positions of the three bottom layers of the slab were kept fixed to the ideal bulk positions, all other atoms were relaxed till forces were lower than 0.005 eV/Å.

5.5 Conclusion

In this work we studied the adsorption of HBC molecules on a Pt(111) surface. The starting point of this study was to find out the adsorption geometry of the molecule. Combining the STM measurements and the DFT simulations it was found out that HBC molecule adsorbs on top site in two symmetry-equivalent configurations. In the low-coverage experiment it was also noticed that the HBC molecules do not form domains. This observation suggested that the intermolecular interactions on a Pt(111) surface are mostly repulsive unlike on other surfaces such as Cu(111) and Au(111).

To find out the origin of the repulsive interactions we focused on the higher coverage of HBC on Pt(111). At this conditions the formation of only non-closely packed (7, 2) and (7, 3) domains and their symmetry analogous (7, 5) and (7, 4) was

observed experimentally, while the formation of the (6, ...) domains was not reported. We addressed the question of the domains' stability to the DFT simulations. The simulations revealed that when HBC molecule adsorbs on the Pt(111) it loses some amount of the electron density that gets transferred to the surface of the metal. The transferred density forms an extended electronic cloud which surrounds the molecule and overlaps in the case of the close-packed (6, ...) domains resulting in the considerable increase of the Pauli repulsion energy. In the case of (7, ...) the transferred charge-density of neighboring molecules barely overlaps, making those domains significantly more stable.

Chapter 6

ProbeParticle model: evaluation and improvement

Noncontact Atomic Force Microscopy (nc-AFM) is a widely used experimental tool to study the surfaces of different materials as well as the molecules adsorbed on them. Following the introduction to the nc-AFM technique given in section 1.5 and the discussion of the application of the most commonly used modeling approaches for the nc-AFM simulations given in section 2.10, we will now focus on our attempts to improve an existing classical model called ProbeParticle (PP) [34, 33]. Despite being very simple (only one particle representing the movement of CO, absence of a metallic tip, classical interactions), PP appeared to be a very promising model.

A clear advantage of the PP model is its speed: it allows to simulate nc-AFM images in a couple of minutes on a single-CPU machine (openCL implementation reduced the time to a couple of seconds performing the calculations on a graphics card), while DFT-based nc-AFM simulations (DFT-AFM) sometimes require a month of calculations on a supercomputer. The drawback of PP is that it is parameter-based and some parameters (such as the charge of PP) are not physically well-defined.

In this work, we tried to take the PP model out of its "comfort zone"[1] and to push it towards its limits by fitting the parameters of the model to the results of DFT-AFM simulations, considering the latter as a "gold standard". The reference DFT tip model consisted of 2 Copper atoms with a CO molecule attached to it – a

[1] By "comfort zone" we mean the type of applications that PP was originally designed for: to provide qualitative agreement with respect to the experimental nc-AFM images.

minimalistic setup which has been successfully employed by Gross et al. [30] and has proven its reliability. This approach allowed to systematically investigate the weaknesses of the PP model and to propose an improvement to it.

We considered five small organic molecules that are drawn in the Figure 6.1. The choice was based on availability of the experimental results that could be used as a reference and on diversity of chemical groups so that the results of the work can be applied to molecules of the same classes.

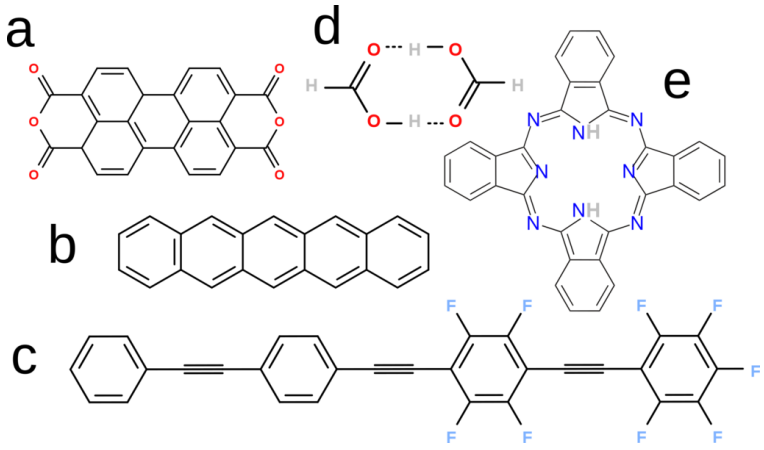


FIGURE 6.1: Five molecules investigated in the current study: (a) PTCDA, (b) pentacene, (c) FFPB, (d) formic acid dimer (FAD), (e) phtalocyanine.

We demonstrated that in order for the PP to be in quantitative agreement with DFT-AFM and to transfer the approach to different systems, the former has to be improved. Therefore, we introduced a new, so-called 2-point-ProbeParticle model (2PP) to also account for the carbon atom of the CO molecule. It turned out that the new model resulted in both: better qualitative and quantitative improvement with respect to standard PP. More specifically, the line-profiles computed with the 2PP model agreed with DFT-AFM in a wider range of the tip's heights, the charges of the particles in the 2PP model become physically well-defined and the fitted parameters of the tip agree better for different molecules.

Based on the results of the DFT simulations with a Cu_2CO tip placed on top of the molecules mentioned above, we have also demonstrated that the tip's CO experiences a polarization which depends on the sample. The importance of accounting for such an effect for a successful nc-AFM model was also highlighted by Kim and

Chelikowsky [163]. To effectively incorporate polarization in the 2PP model, we employed the Restraint Electrostatic Potential (RESP) [110, 164] method to compute the charges of the tip's atoms based on the electrostatic potential obtained with DFT. The resulting RESP charges could then be directly plugged into the 2PP model with no additional modifications.

6.1 nc-AFM models

In Figure 6.2 we show the three tip models used in this study: the standard ProbeParticle model (PP), the DFT model (DFT-AFM) and the 2-point-ProbeParticle (2PP). The PP model has three types of interactions with the sample: vdW attraction as

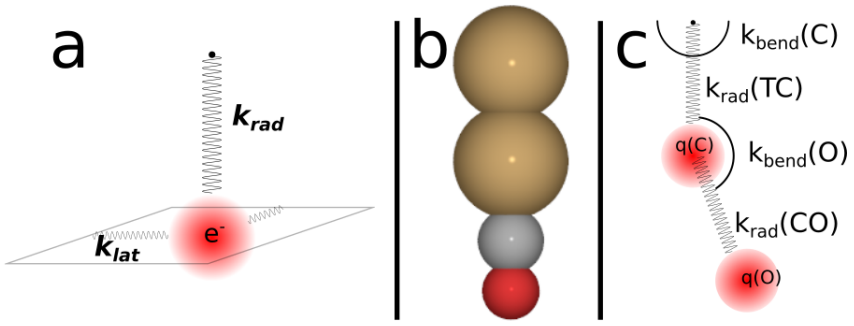


FIGURE 6.2: The three tip models used in the current study: (a) the standard ProbeParticle (PP) model, (b) the DFT-AFM model and (c) the 2-point-ProbeParticle model.

the C_6 term of the Lennard-Jonnes (LJ) potential; Pauli repulsion as the C_{12} term of the LJ potential; and the electrostatic interaction obtained as a convolution of the electrostatic potential of the sample with the spherically symmetric charge density around the ProbeParticle. Therefore, those types of interactions required ε and r_{min} (LJ parameters defined in section 2.10, Equation 2.74) and the total charge Q_{PP} of the particle as parameters to be predefined. To account for the relaxations, PP is attached to the end of a tip[2] through a radial spring with the stiffness k_{rad}, while the lateral displacements are controlled by the lateral springs with the stiffness k_{lat}. In total, the standard PP model of the tip required 5 parameters.

The 2PP model interacts with the sample exactly in the same manner as the standard PP, but has two ProbeParticles instead of one and more LJ parameters: (ε_{lat}(C),

[2]Please note that in the original PP the model tip is represented by a point in space with no interactions with the sample.

$R_{min}(\mathrm{C})$, $\varepsilon_{lat}(\mathrm{O})$, $R_{min}(\mathrm{O})$). To account for the electrostatic interactions between the CO molecule and the sample, we have also included the charge of the Carbon $Q(\mathrm{C})$ and Oxygen $Q(\mathrm{O})$ atoms. To be fully compatible with the DFT model, we introduced 2 copper atoms of the tip: $Q(\mathrm{Cu(up)})$ and $Q(\mathrm{Cu(down)})$ that interact with the sample through electrostatic interactions only (vdW interactions were neglected due to the long distance between the Cu atoms of the tip and the sample). Instead of fitting the charges of the four tips' atoms, we used values derived specifically for each tip-sample combination with the RESP method [110, 164]. The radial stiffness of the Cu-C and C-O bonds was neglected assuming rigid bonds. Lateral stiffness of the PP was replaced with two angular stiffnesses: Cu-Cu-C with the force constant ($k_{lat}(\mathrm{C})$) and Cu-C-O with the force constant ($k_{lat}(\mathrm{O})$). In total, the 2PP model of the tip required 6 parameters to be optimized.

The DFT-AFM tip model (used as a "gold standard") contained 4 atoms: 2 Cu atoms, representing the metallic part of the tip, and a CO molecule. During optimization the CO molecule was allowed to relax, while the Cu dimer was fixed. The z-component of the force acting from the sample on the tip was computed as the algebraic sum of all z-components of forces acting from the sample on every atom of the tip (see the discussion in section 2.10 for more details).

Sample molecules studied in this work were kept rigid and were exactly the same for all the models, allowing a direct comparison among them.

6.2 CO polarization and RESP charges

As it was anticipated earlier, the charges of the tip's atoms depend considerably on the sample under investigation. This dependence can be emphasized with the help of the electronic density difference defined as

$$\rho(\vec{r}) = \rho_{smpl-tip}(\vec{r}) - \rho_{smpl}(\vec{r}) - \rho_{tip}(\vec{r}) \tag{6.1}$$

where $\rho_{smpl-tip}(\vec{r})$ is the charge-density of the tip together with the sample, $\rho_{smpl}(\vec{r})$ is the charge-density of the lone sample and $\rho_{tip}(\vec{r})$ is the charge-density of the lone tip. The $\rho(\vec{r})$ shown in Figure 6.3 reveals the influence of the different samples on

the electronic density distribution. The most noticeable dissimilarity here is the one

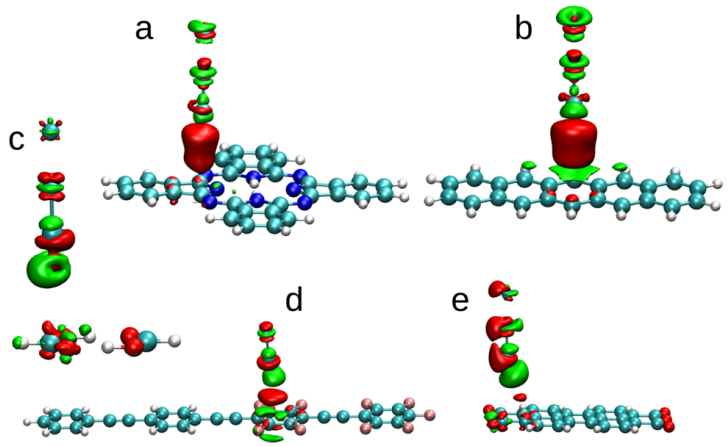


FIGURE 6.3: Electron density distribution obtained for (a) phtalocyanine, (b) pentacene, (c) formic acid dimer, (d) FFPB and (e) PTCDA molecules. The green isosurface represents the region with an excess of electron-density, while the red isosurfaces represent the region with a lack of electron density with respect to the non-interacting situation.

between FAD[3] and PTCDA[4] on the one side (Figure 6.3c,e) and phtalocyanine, pentacene and FFPB[5] one the another (Figure 6.3a,b,d). In the first case, the CO tip is located on top of the carboxylic carbon atom, which, due to the high electronegativity of oxygen, has a partial positive charge. Therefore, it attracts the electronic density of the CO molecule, as represented with a green cloud in front of it (Figure 6.3c,e). When the CO molecule is located on top of negatively charged carbons, it results in a lack of electronic density in front of the former (see Figure 6.3a,b,d).

This situation is well reflected in the RESP charges computed for the tip being on top of different molecules, Table 6.1. For the case of FAD and PTCDA, the charge of the oxygen atom is the most negative (-0.13 e for both cases), while the charge of the carbon atom is the most positive (+0.26 for FAD and +0.23 for PTCDA).

As the polarization of the CO molecule is reflected in the RESP charges, we used those values to effectively include the polarization effect of the CO molecule of the tip being on top of a particular sample.

[3]Formic acid dimer.
[4] Perylenetetracarboxylic dianhydride.
[5]4-(4-(2,3,4,5,6-pentafluorophenylethynyl)-2,3,5,6-tetrafluorophenylethynyl) phenylethynylbenzene.

TABLE 6.1: Average RESP charges (in electrons) computed for the tip being on top of all the molecules investigated in the current study.

	phtalocyanine	pentacene	FAD	FFPB	PTCDA
Cu (up)	-0.06	-0.07	-0.03	-0.06	-0.05
Cu (down)	-0.04	-0.06	-0.10	-0.06	-0.07
C	+0.14	+0.21	+0.26	+0.19	+0.23
O	-0.06	-0.12	-0.13	-0.10	-0.13

6.3 Tip parameters for PP and 2PP models

We started our investigations with the fitting of the PP and 2PP tip models to the DFT-AFM simulations. We chose PTCDA as the reference molecule since it contains carboxylic groups with significant polarization, implying the important role of electrostatic interactions between the tip and sample. The LJ parameters for the molecule were taken from the OPLS-AA force field [165, 166] and were not optimized in this case. The fitted set of parameters are shown below in Table 6.2 for the PP model and in Table 6.3 for the 2PP model.

TABLE 6.2: Parameters of the PP model used in the current study.

Q, [e]	k_{lat}, [N/m]	k_{rad}, [N/m]	r_{min}, [Å]	ε, [eV]
-0.03	0.35	21.9	1.58	1.88×10^{-2}

TABLE 6.3: Parameters of the 2PP model used in the current study.

$k_{lat}(C)$ [eV/rad^2]	$k_{lat}(O)$ [eV/rad^2]	$r_{min}(C)$ [Å]	$\varepsilon(C)$ [eV]	$R_{min}(O)$ [Å]	$\varepsilon(O)$ [eV]
0.25	0.15	1.51	6.2×10^{-2}	1.83	3.5×10^{-3}

After having optimized those tip parameters, we fixed them and did not change them anymore in the course of the current study. The LJ parameters (r_{min} and ε) of the molecules (including PTCDA), however, were optimized.

Further in the text we will go through every molecule and compare the PP/2PP

to DFT-AFM and experimental results (where available). We will focus on the accuracy that both the PP and 2PP models could achieve.

6.4 Case study: PTCDA

PTCDA was studied experimentally and theoretically in the work of Moll et al. [167] (PTCDA on a Cu(111) surface) and Iwata et al. [168] (PTCDA on a Si(111) surface). The experimental nc-AFM image of PTCDA adsorbed on a Si(111) surface is shown in Figure 6.4.

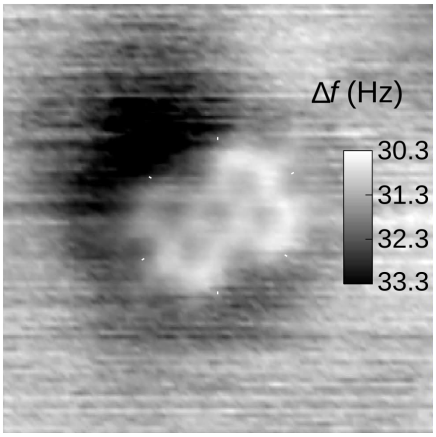
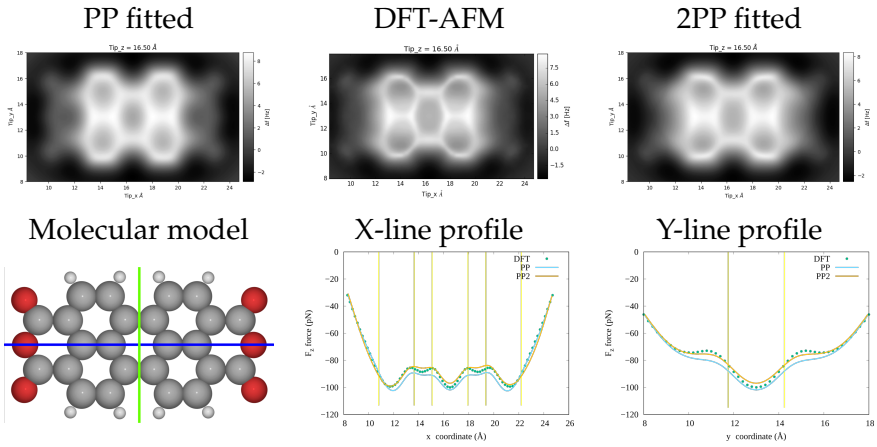


FIGURE 6.4: Constant-height nc-AFM image of a PTCDA molecule adsorbed on a Si(111) surface. The figure is taken and adapted from the reference [168].

Several important features of the experimental nc-AFM images are noticeable. Firstly, one could see that the appearance of the anhydride groups is significantly less pronounced compared to the carbon rings of the molecule. Further, one could notice that all the carbon rings of the molecule are considerably stretched, while the outer rings are additionally distorted. All these features are present and correctly reproduced by the DFT-AFM model, see Table 6.4. Both the PP and 2PP models fitted to the DFT results are able to reproduce those features as well. However, it is evident that the standard PP model is more "rough" in this respect: the outer carbon rings are considerably less tilted compared to the case of the DFT-AFM and 2PP models. Also, the PP model does not reproduce the "anhydride region" very well: one could see the straight lines corresponding to the C-C bond, while in the case of DFT-AFM and 2PP those features are quite rounded. Looking at the line profiles one could understand the origin of better agreement of the 2PP model compared to

the standard PP. The flexibility of the former model allows to better reproduce the curvature of the profiles, which is essential for nc-AFM imaging. The line-profiles of the standard PP model are more straight, which results in higher noise in the images and less sharp bonds. The fitted LJ parameters obtained for the PTCDA molecule can be further reused for the aromatic hydrocarbons and anhydrides of the organic acids.

TABLE 6.4: PTCDA molecule. First row: comparison of the nc-AFM images obtained with the PP, DFT and 2PP models. CO-sample distance is 3.5 Å. Second row: image representing the molecule, where gray spheres correspond to carbon atoms, white spheres correspond to hydrogen atoms and red spheres correspond to oxygen atoms. The X-line profile is the z component of the force acting from the sample on the tip at height 3.5 Å obtained as a function of the position on the blue line. The Y-line profile is similar to the X-line profile but obtained along the green line. The vertical yellow lines correspond to the positions of atoms or bonds on the blue and green lines drawn on the molecule.



6.5 Case study: Pentacene

A pentacene molecule adsorbed on a Cu(111) surface was studied both experimentally and theoretically in the work of Gross et al. [29] and theoretically in the work of Guo et al. [131]. The experimental nc-AFM image of pentacene shows that the edges of the molecule appear significantly brighter (see Figure 6.5).

This feature is a combination of two effects. One is purely geometrical: the pentacene adsorbed on the Cu(111) is distorted in such a way that the edges are slightly lifted up. Therefore, the bonds that are closer to the tip appear significantly brighter.

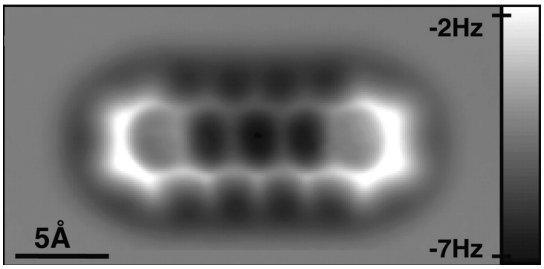
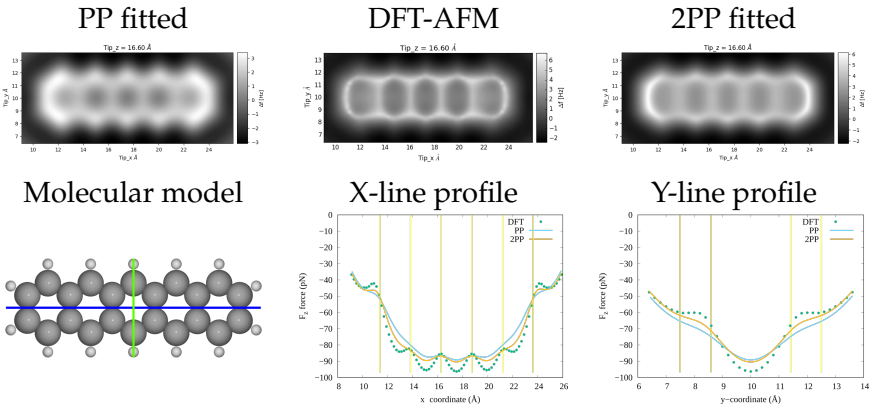


FIGURE 6.5: Constant-height nc-AFM image of a pentacene molecule adsorbed on a Cu(111) surface. The figure is taken and adapted from the reference [29].

The second reason is that the tip, being placed at the edge of the molecule, experiences a lower attraction force due to less interaction with the molecule (see Table 6.5). This feature has been successfully reproduced by all the models used in this study.

TABLE 6.5: Pentacene molecule. First row: comparison of the nc-AFM images obtained with the PP, DFT and 2PP models. CO-sample distance is 3.6 Å. Second row: image representing the molecule, where the gray spheres correspond to carbon atoms, white spheres correspond to hydrogen atoms. The X-line profile is the z component of the force acting from the sample on the tip at height 3.6 Å obtained as a function of the position on the blue line. The Y-line profile is similar to the X-line profile but obtained along the green line. The vertical yellow lines correspond to the positions of atoms or bonds on the blue and green lines drawn on the molecule.



However, the nc-AFM image of a molecule simulated with the PP model appears quite wider in the y-direction with respect to DFT-AFM, while the 2PP model agrees pretty well in this respect. And again, one could see that the PP nc-AFM image is quite "rough" compared to 2PP: the bright edges of the pentacene in the 2PP model are significantly more rounded compared to the standard PP — very similar to the DFT-AFM results and to the experiment. The line profiles obtained with the PP and

2PP models follow the same trend: the PP profiles are also quite rough, while the 2PP profiles are more curved, however not as curved as the DFT-AFM line profiles.

We note here that for the pentacene molecule, the tip parameters were not optimized as the values from the PTCDA were reused. Obviously, one could also optimize the tip parameters for this particular case, achieving better agreement with respect to DFT-AFM, however this would make our approach not expendable in studying a wider set of systems. The tip's parameters (apart from the charges as it was shown earlier) should be substrate independent. On the other hand, the LJ parameters of atoms of the pentacene molecule were optimized, but they can be reused further for a wide class of organic molecules such as aromatic hydrocarbons, graphene nanoribbons, etc.

6.6 Case study: FFPB

An FFPB molecule adsorbed on a two-monolayer thick (100) NaCl film on Cu(111) was studied both experimentally and theoretically in the work of Moll et al. [136]. The nc-AFM image of the molecule (see Figure 6.6) has several noticeable features: the fluorinated (F-rings) and hydrogenated (H-rings) rings appear different in size, as the hydrogenated ones are more extended. Moreover, the H-rings overall appear darker compared to the F-rings. It is also noticeable that all the triple bonds possess a bright protrusions which was observed both theoretically and experimentally [136].

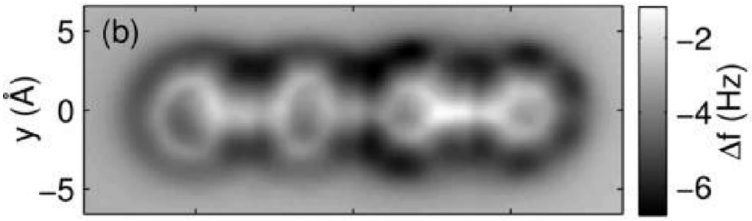
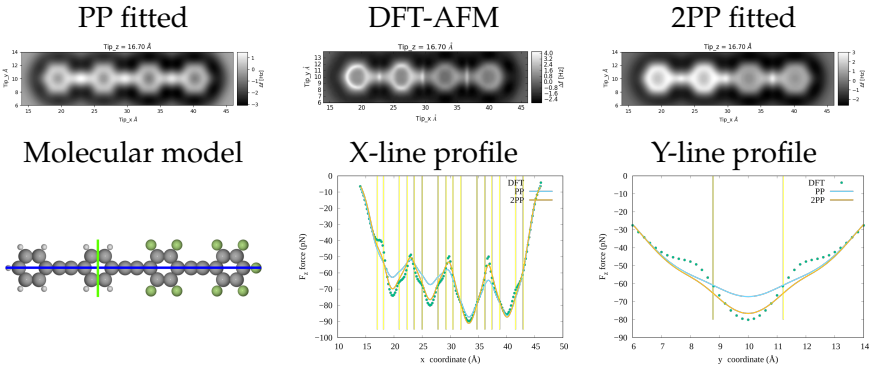


FIGURE 6.6: Constant-height nc-AFM image of an FFPB molecule adsorbed on a two-monolayer thick (100) NaCl film on Cu(111). The figure is taken and adapted from the reference [136].

All the mentioned features are reproduced very well by the DFT-AFM as well as by the 2PP model (see Table 6.6), however some tiny discrepancies remain: the F-rings in the 2PP are more blurred compared to the DFT-AFM simulations and the

bonds in DFT are more rounded compared to both PP models. The standard PP model has two additional disagreements which are: the difference in appearance of the H-rings and F-rings is not so noticeable as they have more or less the same size and the difference in brightness is not so noticeable.

TABLE 6.6: FFPB molecule. First row: comparison of the nc-AFM images obtained with the PP, DFT and 2PP models. CO-sample distance is 3.7 Å. Second row: image representing the molecule , where the gray spheres correspond to carbon atoms, white spheres correspond to hydrogen atoms and green spheres correspond to the fluorine atoms. The X-line profile is the z component of the force acting from the sample on the tip at height 3.7 Å obtained as a function of the position on the blue line. The Y-line profile is similar to the X-line profile but obtained along the green line. The vertical yellow lines correspond to the positions of atoms or bonds on the blue and green lines drawn on the molecule.



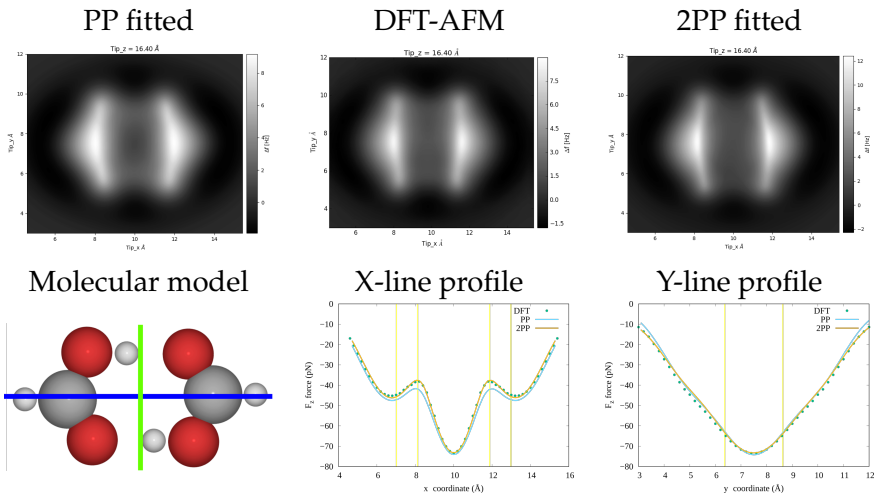
The line profiles of the 2PP model are in very good agreement with the DFT-AFM line profiles, while the standard PP model is still lacking flexibility and cannot adapt very well to the DFT-AFM results. However, both models fail to correctly reproduce the Y-profile at 8-9 and 11-12 Å. The LJ parameters obtained for the FFPB molecule can be further reused for other fluorinated aromatic compounds and compounds containing a triple carbon-carbon bond.

6.7 Case study: formic acid dimer

To the best of our knowledge, the experimental nc-AFM images of the formic acid dimer (FAD) or its analogs were not published, therefore we focus our attention only on the agreement of PP and 2PP with the DFT-AFM results. Interestingly, looking at

the results of both empirical models one notices excellent agreement with respect to DFT-AFM in terms of the nc-AFM images and line profiles (see Table 6.7).

TABLE 6.7: Formic acid dimer. First row: comparison of the nc-AFM images obtained with the PP, DFT and 2PP models. CO-sample distance is 3.4 Å. Second row: image representing the molecule, where the gray spheres correspond to carbon atoms, white spheres correspond to hydrogen atoms and red spheres correspond to oxygen atoms. The X-line profile is the z component of the force acting from the sample on the tip at height 3.4 Å obtained as a function of the position on the blue line. The Y-line profile is similar to the X-line profile but obtained along the green line. The vertical yellow lines correspond to the positions of atoms or bonds on the blue and green lines drawn on the molecule.



Such an excellent agreement is explained by two factors. The first factor (and the most important one) is the very smooth line profile that is easily reproduced by the PP and 2PP models. The second factor has to do with the tip parameters. The PTCDA molecule used to make such a fit chemically corresponds pretty well to the formic acid dimer. The LJ parameters obtained for the formic acid dimer can be further reused for other organic acids.

6.8 Case study: phtalocyanine

We investigated the phtalocyanine molecule, an analog of naphtalocyanine. The latter adsorbed on a 2 monolayer thin insulating film of NaCl on a Cu(111) surface was studied experimentally and theoretically in the work of Mohn et al. [31]. The experimental nc-AFM image (see Figure 6.7) reveals that the inner part of the molecule is

darker compared to the outer carbon rings. Unlike in STM, the asymmetry of the system at the level of the nitrogen atoms (N-lobes and H-lobes) is not pronounced [31]. The DFT-AFM model, as well as both PP and 2PP, also predict the inner region to

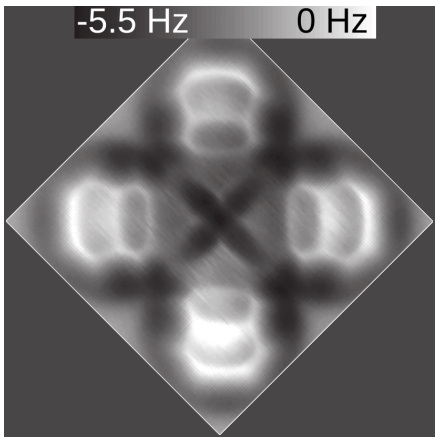


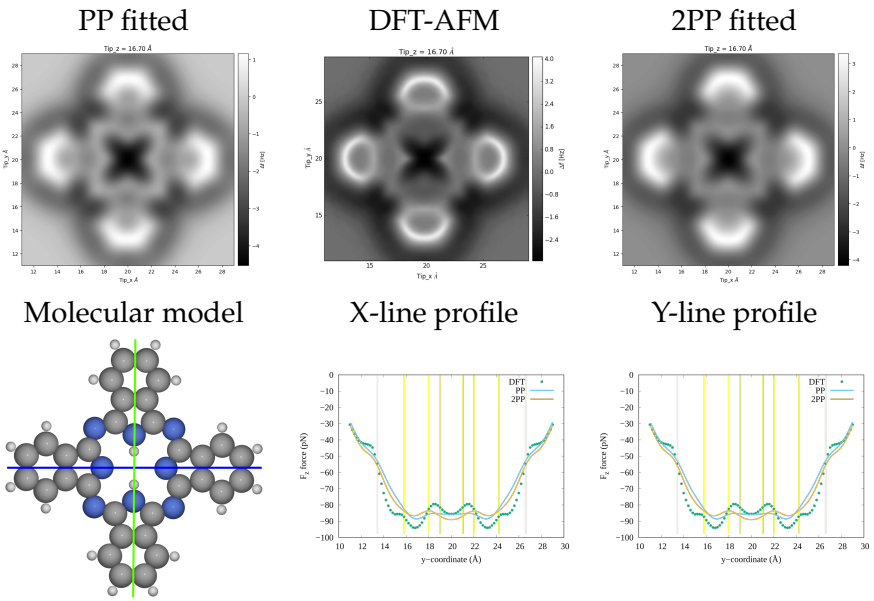
FIGURE 6.7: Constant-height nc-AFM image of the naphtalocyanine molecule adsorbed on a 2 monolayer thin insulating film of NaCl on Cu(111). The figure is taken and adapted from the reference [31].

be less bright compared to the outer carbon rings. The C_4N ring is smeared out in the DFT-AFM and 2PP simulations, however in the standard PP model bonds are visible: from this perspective, the 2PP model performs slightly better. In the current example, both the PP and 2PP models are unable to reproduce the roundness of the outer carbon ring, which is predicted by the DFT-AFM model and visible in the experiment. The cross inside of the inner part of the molecule is reproduced quite well by DFT-AFM, while PP and 2PP perform quite poorly here. Furthermore, the shape of the inner part - which is quite round in both DFT-AFM and the experiment in the PP and 2PP models - becomes rather quadratic.

The line profiles agree quite poorly with the DFT simulations. We again notice here that the line profiles with quite a complex shape are still difficult to be reproduced by the PP and 2PP models. To improve the performance of the models for such cases, we can see two possibilities. The first would be to replace the LJ potential responsible for Pauli repulsion and vdW attraction with some other potential with an adjustable width. An example of such a potential can be the Morse potential $V(r) = D_e \left(1 - e^{-a(r-r_e)}\right)^2$, where r is the distance between the atoms, r_e is the equilibrium distance, D_e is the depth of the potential and a controls the "width" of the

potential. This would allow to obtain the potential energy surface of a more complex shape and potentially solve the problem of the curvature of the line profiles. The second possibility would be to explore the polarizable fore fields [169] to see if they could also help in reproducing the curvature of the line profiles and provide better nc-AFM images.

TABLE 6.8: Phtalocyanine molecule. First row: comparison of the nc-AFM images obtained with the PP, DFT and 2PP models. CO-sample distance is 3.7 Å. Second row: image representing the molecule, where the gray spheres correspond to carbon atoms, white spheres correspond to hydrogen atoms and blue spheres correspond to nitrogen atoms. The X-line profile isthe z component of the force acting from the sample on the tip at height 3.7 Å obtained as a function of the position on the blue line. The Y-line profile is similar to the X-line profile but obtained along the green line. The vertical yellow lines correspond to the positions of atoms or bonds on the blue and green lines drawn on the molecule.



6.9 Fitted Lennard-Jones parameters

Having discussed the agreement of PP and 2PP simulations *vs* DFT-AFM and experimental results we now focus on the comparison of the fitted Lennard-Jones atomic parameters. The parameters are all summarized in Table 6.9 at the end of this chapter. For the reader's convenience, parts of the molecules with highlighted atomic types are also provided.

It can be seen that in the case of 2PP model the fitting procecure performed pretty well. The fitted r_{min} values for the same atomic types but different molecules remained very close to each other. Contrary to this, the PP model provided quite scattered r_{min} values. A clear example to support this is the C_{carb} atom type, where in the case of PP the difference between PTCDA and FAD is 18%, while for the 2PP model it is $\approx 1\%$. The maximum deviation for the 2PP models is is less then 4%.

Interesingly, both models provide substantially different r_{min} values for oxygen atoms in the case of PTCDA and FAD molecules. The difference is especially pronunced for the case of O_{sp3} oxygen. We attribute this difference to the fact, that in the case of FAD the O_{sp3} atom is connected to C_{carb} and H, while in the case of PTCDA the oxygen is surroundded by two C_{carb} atoms. As the result the O_{sp3}/PTCDA has less electron density which results in the smaller r_{min} value.

While for the case of r_{min} values fitting 2PP model was significantly better then the standard PP model – for the case of ε fitting both models performed quite poorly. If we consider all atoms except hydrogen[6] – the ε values varried significantly for both models. However, the ratio between the smallest and the largest values was always < 2. For the case of hydrogen atoms both models performed very bad as the ε values vary by an order of magnitude.

Thefore, we can conclude that the nc-AFM is sensitive to the r_{min} parameters and significantly less sensitive to the ε ones. This makes very difficult the task of obtaining the ε parameters from nc-AFM experiment. However, this also reveals that the results of the PP/2PP are not very sensitive to the quality of the ε parameters.

6.10 Methods

The geometries of all the molecules were optimized using the CP2K package [97, 96]. The PBE [54]-D3 [69] approximation to the DFT was used through the whole study. The coordinates of the molecules were initially optimized and then kept fixed for all the subsequent nc-AFM simulations. The Gaussian and augmented plane-waves method [170, 171] was used to solve the electronic equation [98]. DZVP basis sets were employed for copper atoms and DZVP basis sets [156] for all the others. For

[6] We do not consider oxygen atoms here, as, acccording to the discussion above, their chemical environment is well different.

the geometry optimizations the cutoff was set to 400 Ry, while for the subsequent force calculations with fixed geometry the cutoff was increased to 800 Ry. The core-electrons were represented with Goedecker-Teter-Hutter pseudopotentials [157].

Since the geometries provided for DFT-AFM and for the PP models were exactly the same, it was possible to directly compare the vertical forces obtained with both approaches. The following sum of the square differences was applied as a figure-of-merit: $\sigma^2 = \sum_i^{grid}(F_z^i(DFT) - F_z^i(PP))^2$, where $F_z^i(DFT)$ is the total vertical force acting on the tip computed within the DFT-AFM model and $F_z^i(PP)$ is the one computed within the PP models.

The Nelder-Mead method [172, 173] implemented within the Scipy package [174, 175][7] was used to minimize the σ^2. The simulated nc-AFM images were plotted employing the Matplotlib package [178]. The LJ parameters from the OPLS-AA force-field [165, 166] were used as the starting values for the optimization procedure.

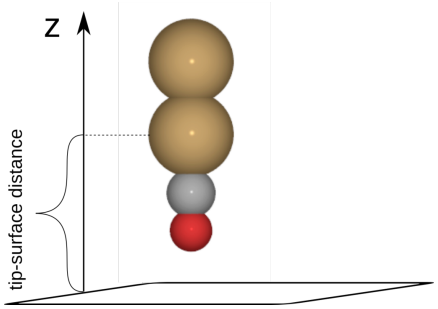


FIGURE 6.8: Definition of the tip-substrate height used in the current study.

For both the PP and 2PP models, the tip parameters were optimized using the PTCDA molecule with OPLS-AA [165, 166] as the sample. After that the parameters of the PP and 2PP models were kept fixed, but the LJ parameters of all the molecules were refitted. The set of the molecules was selected in such a way that the results could be compared to already published experimental results: pentacene [29], formic acid dimer, phthalocyanine [31], FFPB [136] and PTCDA [167, 168].

Each molecule in this work was investigated by DFT-AFM, where the tip was composed of a Cu dimer with the CO molecule attached to it Figure 6.8. The scanning grid covered the asymmetric part of molecules with a density of one grid point

[7] "...open source scientific tools for Python [176, 177]"

per 0.2 Å in the lateral directions and one grid point per 0.1 Å in the vertical direction. The tip was approached from 7.0 Å to 6.2 Å (for the tip height definition see Figure 6.8) above the surface.

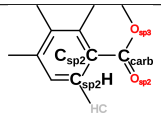
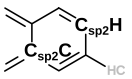
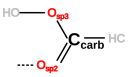
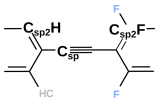
6.11 Conclusion and outlook

In this work, we fitted the PP and 2PP models to the vertical forces of DFT-AFM simulations of a PTCDA molecule. Comparing the results of both classical models, we noticed that introducing the second ProbeParticle allows to improve the quantitative and qualitative agreement with respect to the DFT-AFM results. The 2PP model works better for several reasons. Firstly, two ProbeParticles provide more flexibility and, therefore, allow to reproduce more complex DFT-line profiles. Secondly, the dipole moment of the CO molecule is quite well reproduced by the 2 particles, which provides additional enhancement of the simulated results reproducing better electrostatic interactions with the sample molecule. However, looking at the results of the nc-AFM simulations of the phtalocyanine molecule (see section 6.8), we notice that the 2PP model still has room for improvement. We anticipate that the remaining discrepancies between DFT-AFM and 2PP can be resolved by introducing either a more tunable potential such as the Morse potential, or introducing a polarizable force field to correct for the dynamic electron-density response of the CO molecule to the sample.

In this study we did not explore the influence of the substrate on the adsorbed molecule. The presence of the former can deform the shape of the molecule or can modify its atomic vdW parameters. The first case is already covered by the current approach. It only requires an additional calculation to perform the geometry optimization of the molecule on the substrate. The presence of the substrate (at least in the simplest symmetric cases of adsorption on flat close-packed surface) will only add a background force that can be barely detected in nc-AFM measurements. The latter case (chemisorption), in contrast, seems to be quite complex and requires additional investigations. In this case, the LJ parameters of the molecule would need to be adapted for a particular substrate, making our approach unfeasible. Investigation

of this topic would be an interesting task, however it goes beyond the scope of the current study.

TABLE 6.9: Fitted Lennard-Jones parameters obtained for PTCDA, pentacene, FAD, FFPB and phtalocyanine molecules. For simplicity, the fragments of the molecules are provided highlighting different atomic types. Every cell contains r_{min} (Å, first number) and ε (eV, second number) values for both, PP and 2PP models.

Atomic type	PTCDA	Pentacene	FAD	FFPB	Phtalocyanine
					
$C_{sp2}C$	**PP:** 1.95; 3.31×10^{-3} **2PP:** 1.96; 2.96×10^{-3}	**PP:** 1.81; 6.15×10^{-3} **2PP:** 1.89; 4.90×10^{-3}			**PP:** 1.87; 4.05×10^{-3} **2PP:** 1.92; 1.61×10^{-3}
$C_{sp2}H$	**PP:** 2.11; 1.75×10^{-3} **2PP:** 1.99; 3.62×10^{-3}	**PP:** 2.13; 1.06×10^{-3} **2PP:** 2.03; 3.14×10^{-3}		**PP:** 1.99; 2.04×10^{-3} **2PP:** 1.95; 5.69×10^{-3}	**PP:** 2.08; 1.26×10^{-3} **2PP:** 1.99; 4.57×10^{-3}
C_{carb}	**PP:** 1.92; 1.36×10^{-3} **2PP:** 2.00; 1.96×10^{-3}		**PP:** 2.14; 2.15×10^{-4} **2PP:** 1.98; 1.20×10^{-3}		
$C_{sp2}F$				**PP:** 1.96; 3.66×10^{-3} **2PP:** 1.90; 6.03×10^{-3}	
C_{sp}				**PP:** 1.93; 4.05×10^{-3} **2PP:** 2.08; 1.48×10^{-3}	
O_{sp3}	**PP:** 1.72; 1.38×10^{-2} **2PP:** 1.66; 4.25×10^{-3}		**PP:** 1.98; 2.47×10^{-3} **2PP:** 1.82; 6.26×10^{-3}		
O_{sp2}	**PP:** 1.76; 1.18×10^{-2} **2PP:** 1.65; 1.19×10^{-2}		**PP:** 1.92; 3.80×10^{-3} **2PP:** 1.77; 9.18×10^{-3}		
N					**PP:** 1.91; 4.24×10^{-3} **2PP:** 1.79; 7.12×10^{-3}
F				**PP:** 1.54; 8.08×10^{-3} **2PP:** 1.44; 5.80×10^{-3}	
HC	**PP:** 1.33; 4.57×10^{-3} **2PP:** 1.35; 1.30×10^{-3}	**PP:** 1.43; 8.89×10^{-3} **2PP:** 1.39; 1.00×10^{-3}	**PP:** 1.49; 1.52×10^{-2} **2PP:** 1.35; 8.62×10^{-3}	**PP:** 1.53; 5.34×10^{-3} **2PP:** 1.35; 5.04×10^{-4}	**PP:** 1.51; 6.33×10^{-3} **2PP:** 1.35; 2.21×10^{-3}
HO			**PP:** 1.30; 9.77×10^{-3} **2PP:** 1.36; 3.71×10^{-4}		
HN					**PP:** 1.37; 6.67×10^{-4} **2PP:** 1.38; 1.00×10^{-6}

Chapter 7

AiiDA

As it was discussed in the introductory section 1.6, AiiDA (Automated Interactive Infrastructure and Database for Computational Science) is a tool that allows to automate, manage and disseminate calculation data and workflows. By its design, AiiDA reflects the views on how modern computational science should be planned, organized and realized in practice. In this chapter we will describe in greater detail the concepts behind the AiiDA package based on the manuscript by Pizzi and coworkers [45]. We will start with an overview of the Automation, Data, Environment, Sharing (ADES) model that incorporates the conceptual design of AiiDA.

Further in the text we will shift our focus we will focus to the application of AiiDA to modern day computational science. One of the important features of AiiDA is providing a general interface to (in principle) any computational software. However, to link AiiDA to the specific code an AiiDA-plugin should be provided. As in our computational unit of nanotech@surface laboratory at Empa a significant amount of calculations is done using the CP2K package [97, 96], we developed a freely available AiiDA-plugin [179] for CP2K.

We developed an AiiDA-workflow to efficiently tackle the daily operations of computing the electronic properties of 1D materials. The workflow allowed to fully automate the routine operations and to minimize the work of a scientist in preparing the initial geometries of a material.

Another advantage of AiiDA is "coupling to the data", which essentially means that before any data will be used by AiiDA they (and the causal links between them)

will be stored in a database. The practical advantage of such a design is that the user has access not only to the results of the calculations but also to the inputs used to generate them. An example of such a history graph is presented below.

7.1 ADES model

The ADES model incorporates the fundamental ideas behind the AiiDA package. It is based on the four pillars sketched in Figure 7.1 and discussed further.

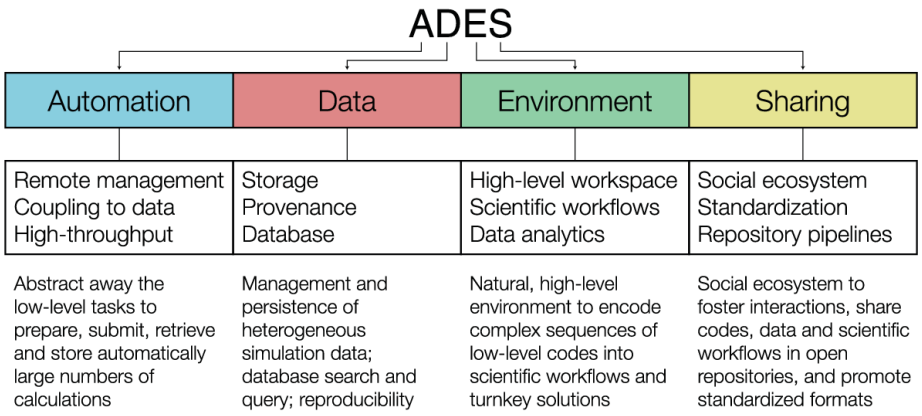


FIGURE 7.1: The ADES model, a fundamental concept of the AiiDA project divided in four main pillars: Automation, Data, Environment and Sharing. The figure is taken from [45].

The first pillar is **automation**. The standard procedure to compute materials properties relies on access to supercomputing facilities. From the design of the calculation to the archiving of the result, a number of collateral operations are needed that can be defined as *remote management*: data have to be copied to the remote machine, the calculations have to be submitted, the status of a calculation has to be checked and the results have to be retrieved. Remote management is based on low-level operations that are basically independent from the specific simulation software and that should be automatized. To remote management belongs also *coupling to data*, i.e. no operations can be performed on data that are not yet stored in the database and the *high-throughput*, i.e. the ability to administrate the execution of a large number of similar calculations with automatic handling of errors.

The second pillar is **data**. While managing the data produced by a calculation, one first decides whether they are important and have to be *stored* or they can be dismissed. This decisional step depends on the specific application and code, thus

a code-dependent policy has to be defined. *Provenance* of the data must be kept in order to provide a path to the original settings which were used to generate the output data. To grant a high efficiency of querying, *databases* should be employed for the storage and organization of the information.

The third pillar is **environment**. A natural environment that complies with the other requirements mentioned before has to be provided to a computational scientist. This environment should include a *high-level workspace* that hides the details of data management. It should support *scientific workflows* that embed scientific knowledge on the correct procedures to simulate specific properties of materials. It should also allow automation of *data analysis* based on the standardization of data types.

The final fourth pillar is **sharing**. Easy sharing of the data enables and promotes interactions among scientists. This can be achieved through the creation of a *social ecosystem* that ensures the privacy of data or allows to share them only when it is needed. Sharing requires *standardization* of the formats and *repository pipelines* to sustain automatic data exchange among different repositories.

7.2 AiiDA – a Python implementation of the ADES model

At this point it is useful to discuss the details on how AiiDA implements the ADES model. There are three main types of objects within AiiDA: **Calculations** which provides a general interface to the diverse compute codes; **Data** which implements particular data storage schemes; and **Computers** that grant standardized access to different computing facilities. A unified protocol to access codes, computers and data is key to automation in submitting calculations, monitoring tasks and retrieving computed results.

The **AiiDA daemon** is a process that follows the run cycle of every task. If a new task is created, the daemon sends inputs to a remote server and submits the calculation to its queuing system. When a calculation is completed, the daemon retrieves the results back to the local machine and parses the computed quantities. From the moment of submission everything happens automatically, no user intervention is required.

Remote management is achieved though plugins which are able to convert generic job attributes (such as maximum execution time, number of CPUs, etc ...) in the submission script specific to the remote scheduler.

The data within AiiDA are organized in a directed acyclic graph[1]. Two data objects can not be directly connected in the graph, a modification is required to transform the "input data" into the "derived data". Physically the graph is stored in a database, which has a complex structure. Only the basic tables[2] are shown here in order to provide the reader with an idea on how object storage is achieved within the AiiDA framework. For more detailed information, the reader is referred to the original manuscript [45] or to the documents section on the AiiDA web-page [45]. First of all, every node (a node is a unit of information stored in the AiiDA graph) has to have a reference number – a unique identifier within the database. Also, being in the graph – each node may have links to input and output nodes, so the links should be stored as well. Different types of objects may have an arbitrary number of attributes, therefore a way for expandable storage of those parameters is required. To achieve all those requirements, the database includes the three tables that are shown in Figure 7.2

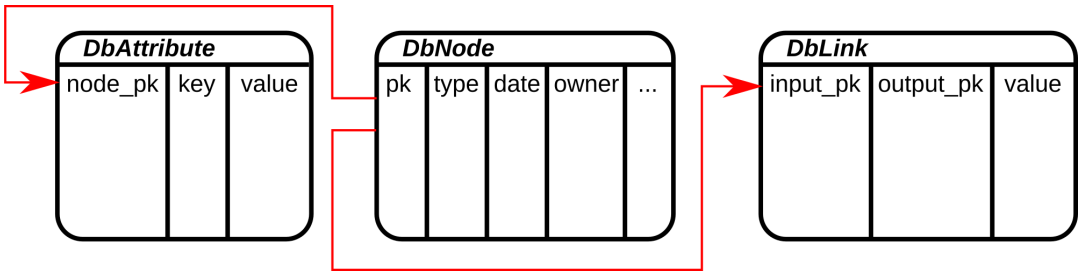


FIGURE 7.2: A schematic representation of the most relevant AiiDA databases. **DbNode** is a database used to store nodes. **DbLink** stores the direct connections between nodes. **DbAttribute** stores nodes attributes (one entry per attribute).

The main table is called **DbNode**. It contains basic information about a particular node: the primary key (pk, a unique identifier of a node), the node type, the creation date, the owner and so on. The links between nodes are stored in the **DbLink** table, where each link is represented by the pk of an input node, the pk of an output node

[1] The links between nodes always have a direction. Loops are not allowed, otherwise it would break the causal relationships between the nodes.

[2] AiiDA has 33 tables defined to manage different attributes, but for simplicity here we refer only to the most important ones.

and a text entry to distinguish the inputs to the same node. An arbitrary number of parameters for every node can be stored having a separate entry for every parameter in the **DbAttribute** table, which has three columns: node pk, parameter key and parameter value. The value of each parameter can be of any type: number, string, matrix, a set of coordinates and so on.

Once a node has been stored, its attributes can not be modified anymore. In addition to entries in the database, AiiDA also allows to store data directly in the file system as needed in the case of large files, such as the charge density of a system defined on a 3D grid, or molecular dynamics trajectories. In that case, the database only contains a reference to the file.

There are three main access points for a user to interact with AiiDA: the command line utility called *verdi*, Python scripts and Python shell. The command line *verdi* is typically used to perform common operations like starting/stopping the daemon, checking the status of calculations, installing/configuring computers and codes. Python scripts and Python shells provide full access to the AiiDA Python application programming interface.

The set of calculations needed to compute a specific material property or several properties is grouped together in a so-called workflow that can be shared among scientists for validation, simple usage and further improvement or development.

Queries to the database to retrieve data can be done in AiiDA via a Python class called **QueryTool**. This class hides the complexity of the database-specific language and provides a user-friendly interface to search/show/analyze stored data.

Standardization of formats simplifies sharing of the results through a combination of databases into one public repository. To rationalize the sharing procedure, AiiDA possesses import/export functionality within the database. To ensure uniqueness, every node holds a Universally Unique IDentifier (UUID) assigned to it. Therefore, if a part of one database was merged into another database, all pk numbers of the first would be automatically updated to be coherent with a new database, while the UUID of each node would be used as a unique node identification. To implement concretely the sharing capabilities of the AiiDA framework, a platform called Materials Cloud [180] has been developed and recently released providing access to

the database of materials properties, workflows, and plugins.

7.3 AiiDA in the nanotech@surfaces laboratory

Most of the DFT simulations performed in the nanotech@surfaces laboratory to compute the gas phase properties of one-dimensional carbon-based nanostructures, or the geometrical and electronic properties of molecules adsorbed on metallic substrates, rely on the CP2K [97, 96] and QUANTUM ESPRESSO [89] compute codes. Highly optimized simulation protocols have been developed in recent years within the lab to efficiently simulate microscopy and spectroscopy properties for a direct comparison with experiments. Band structure calculations, STM images, adsorption geometries and energies are routine calculations, so they can profit considerably from automation protocols based on the AiiDA framework so far discussed. This is why we developed a plugin to interface AiiDA with the most common tasks of the laboratory and a workflow that, starting from the geometry of 1D nanomaterial, computes and archives the most relevant electronic and geometry properties that should be promptly accessible to experimentalists to support them in the design of novel nanostructures. The idea is to provide a web-based platform that should be directly accessible by experimentalists. The work initiated by me was then continued by Ole Schütt and more recently Edward Ditler, culminating in the first release of a web-based AiiDA interface specifically designed for the needs of the laboratory. The work was possible thanks to collaboration with AiiDA developers within the Swiss National Competence Center MARVEL [181].

7.4 CP2K plugin for AiiDA

The CP2K-AiiDA plugin [179] has a simple but general design that allows an interface for a variety of CP2K based calculations. We describe in the following how the input and output plugins work by providing an example of the AiiDA script for performing wavefunction optimization with CP2K.

The script starts with the inclusion of necessary modules (lines 1-9). A molecular builder from the ASE package [182] is loaded (line 3) to allow the creation of input

geometries. Then, all the necessary AiiDA classes are loaded (lines 7-9): **Strucutre-Data** to create a structure for the calculations, **ParameterData** for the computation parameters and **SinglefileData** to attach additional input files.

```
1 from __future__ import print_function
2 import sys
3 import ase.build
4
5 # Import a function to test the code setup as well as basic AiiDA classes
6 from aiida.common.example_helpers import test_and_get_code
7 from aiida.orm.data.structure import StructureData
8 from aiida.orm.data.parameter import ParameterData
9 from aiida.orm.data.singlefile import SinglefileData
```

After this, the information that is required to run the code is obtained (line 16) and a calculation object is created (line 21).

```
10 # Check whether the code is available
11 if len(sys.argv) != 2:
12     print("Usage: test_dft.py <code_name>")
13     sys.exit(1)
14
15 codename = sys.argv[1]
16 code = test_and_get_code(codename, expected_code_type='cp2k')
17
18 print("Testing CP2K ENERGY on H2O (DFT)...")
19
20 # Creating a calculation object
21 calc = code.new_calc()
```

Further, a test system containing a single water molecule is built by means of the ASE molecular builder (lines 23-24). The ASE molecule object further is transformed into an AiiDA StcutureData object (line 25).

```
22 # Building a structure (contains one water molecule only)
23 atoms = ase.build.molecule('H2O')
24 atoms.center(vacuum=2.0)
25 structure = StructureData(ase=atoms)
```

The next step is the choice of parameters for the calculation such as the basis sets, and the pseudopotentials that are provided as a Python dictionary and transformed into a ParameterData object (lines 26-61).

```
26 # Setting up calculation parameters
```

```
27 parameters = ParameterData(dict={
28     'FORCE_EVAL': {
29         'METHOD': 'Quickstep',
30         'DFT': {
31             'BASIS_SET_FILE_NAME': 'BASIS_MOLOPT',
32             'QS': {
33                 'EPS_DEFAULT': 1.0e-12,
34                 'WF_INTERPOLATION': 'ps',
35                 'EXTRAPOLATION_ORDER': 3,
36             },
37             'MGRID': {
38                 'NGRIDS': 4,
39                 'CUTOFF': 280,
40                 'REL_CUTOFF': 30,
41             },
42             'XC': {
43                 'XC_FUNCTIONAL': {
44                     '_': 'LDA',
45                 },
46             },
47             'POISSON': {
48                 'PERIODIC': 'none',
49                 'PSOLVER': 'MT',
50             },
51         },
52         'SUBSYS': {
53             'KIND': [
54                 {'_': 'O', 'BASIS_SET': 'DZVP-MOLOPT-SR-GTH',
55                     'POTENTIAL': 'GTH-LDA-q6'},
56                 {'_': 'H', 'BASIS_SET': 'DZVP-MOLOPT-SR-GTH',
57                     'POTENTIAL': 'GTH-LDA-q1'},
58             ],
59         },
60     }
61 })
```

Once all the inputs are created, they are linked to the calculation object (line 63 – the input structure and line 66 – the input parameters). The input plugin is also able to carry on additional input files that may be required for the calculations (such as pseudopotentials, force field parameters, etc..) employing a **SinglefileData** object, however for this particular case this feature was not utilized.

```
62 # Linking the structure to the calculation object
63 calc.use_structure(structure)
```

```
64
65 # Linking the parameters to the calculation object
66 calc.use_parameters(parameters)
67
68 # Specifying the necessary computational resources
69 calc.set_max_wallclock_seconds(3*60)  # 3 min
70 calc.set_resources({"num_machines": 1})
```

After all the necessary objects are linked to the code and all the necessary computational resources are specified (lines 69-70), all the data are stored in the database (line 72) and the calculation is submitted (line 73).

```
71 # Storing all the nodes in the AiiDA database and submitting the calculation
72 calc.store_all()
73 calc.submit()
```

Once the calculation is completed, the AiiDA daemon downloads the main output file and the trajectory (if available) only. Further, the output part of the CP2K plugin parses the final energy of the system and the coordinates produced at the very last step of geometry optimization or of a molecular dynamics simulation. In the case that specific calculations would require additional files to be retrieved and then parsed, it is possible to specify them with a few additional code lines:

```
settings_dict += {'additional_retrieve_list': ["your_file_pattern"]}
settings = ParameterData(dict=settings_dict)
calc.use_settings(settings)
```

Links to the retrieved files are stored in the database, for later access.

The structure defined above can be easily generalized and used for different kins of calculations. The CP2K-AiiDA plugin is currently being further developed by Ole Shütt and Edward Ditler.

7.5 Workflow to study the electronic properties of 1D materials

In this last section, we provide an example on how to use the developed workflow to compute the gas phase properties of graphene based nanoribbons. Starting from a geometry file describing the periodic unit of a nanoribbon, the first step is to optimize the unit cell and the geometry of the ribbon with a coarse mesh of k-points for

the sampling of the Brillouin zone. After the calculation is converged, the output geometry and cell is used for further optimization done with a more accurate sampling in k space. The number of k-points used in each calculation is automatically decided as a function of the cell size starting from the predefined value obtained from convergence tests that were performed before the creation of the workflow. After a converged cell and geometry has been obtained for the ribbon, the Hartree potential is determined and band structures are computed and aligned to the vacuum level. Finally, spectroscopic properties such as scanning tunneling spectroscopy plots, plots of absolute magnetization and plots of KS orbitals are produced and archived.

An example of results obtained from the workflow is shown in Figure 7.3. A nanoribbon structure with (pk=181) is used as input for cell optimization. The calculation uses additional input objects: parameters (pk=179), set of k-points (pk=180) and the pw.x code[3] (pk=1). The optimized structure (pk=248) is used again for better convergence. After cell optimization is completed, the final structure (pk=333) is further transfered to the geometry optimization performed with a more accurate grid of k-points (pk=346). When the geometry is converged, it is used to compute the wavefunctions (pk=470) with a even denser grid of k-points (pk=468). Finally, the band structure is calculated (pk=881). Having such an automated workflow allows to perform high-throughput calculations to study the electronic properties of a large number of nanoribbons. The results of the scan can be later queried, as shown in Figure 7.4.

7.6 Conclusion

The current level of technologies and the amount of scientific data have already produced calls for an efficient infrastructure to store, share, manage the results and automatize the routine procedures of modern computational science. An attempt to provide such an infrastructure comes with the so-called AiiDA project which is at the core of the NCCR MARVEL initiative. AiiDA provides a general interface for scientific calculations, allowing to prepare them, automatically launch them on

[3]This program is part of the QUANTUM ESPRESSO package

a supercomputer, retrieve the results and store the results in a database that can be further efficiently queried.

To explore the usefulness of AiiDA, we employed it to automate several routine tasks in the computational unit of the nanotech@surface laboratory at Empa. The work started with writing an AiiDA plugin for the CP2K – a quantum chemistry and solid state physics software package that is widely employed in the daily work of the laboratory. As a result, we could achieve almost full automation of such tasks as cell optimization, geometry optimization, molecular dynamics simulations. The automation comes from the fact that working with different materials requires no or very little changes of the AiiDA scripts that manage such jobs.

Further, we explored the automation of calculation chains (or workflows) that is provided by AiiDA. An automated workflow was created that carries a structure through several optimization cycles with gradually increased accuracy and finally computes the electronic and spectroscopic properties of the material. Thanks to the Python-based user-interface, the results of such calculations can be further efficiently analyzed.

In my view, future students working on projects similar to the one that I have faced during my PhD will profit from my work of implementation of workflow and plugins. The automation of the tasks of computation and organization of the results will allow them to reduce the time dedicated to technical tasks and to focus even more than usual on the scientific insight and critical analysis of the results.

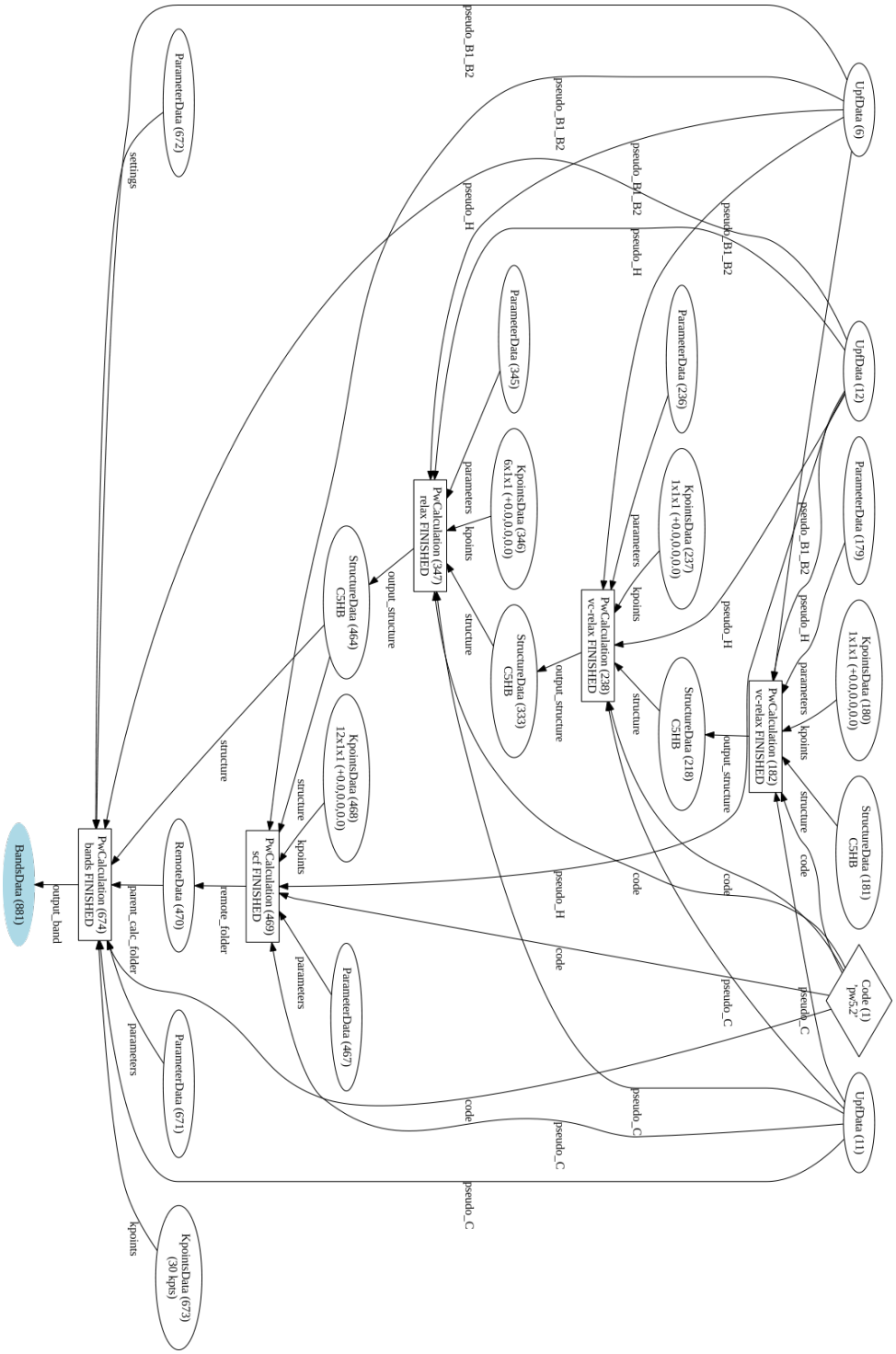


FIGURE 7.3: An example of an AiiDA graph obtained after all the steps of the workflow have been executed.

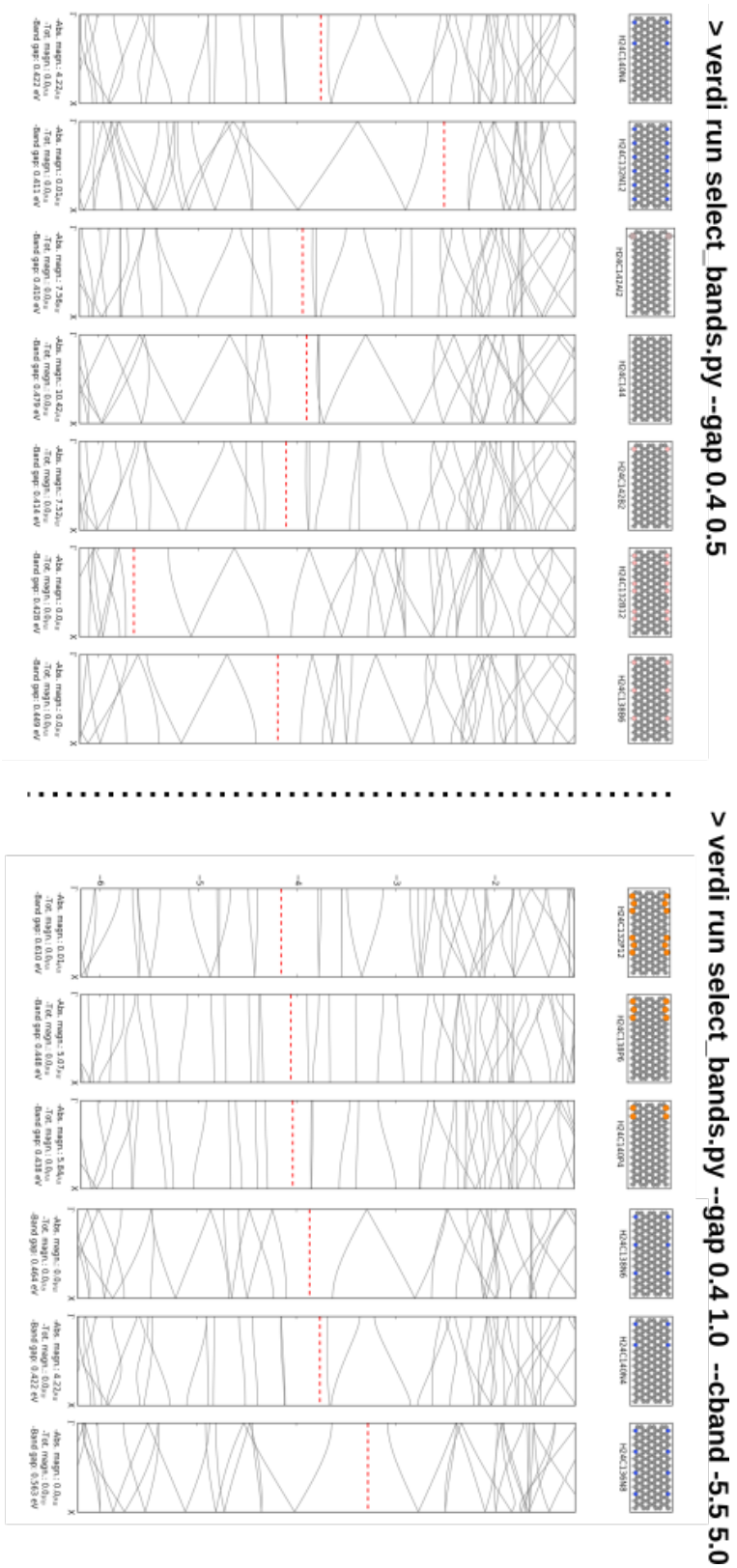


FIGURE 7.4: Two examples of the quarries on the AiiDA database.

Conclusion and outlook

Computational science has gained an importance in offering insight into the microscopic mechanisms underlying observed physical and chemical phenomena. Based on state-of-the-art modeling techniques this work aimed to understand the behavior of adsorbed molecular layers at the surface of metallic and intermetallic systems. This research was done in close contact with experimental groups and was focused both on methodological aspects of computer simulation and on realistic applications to specific cases.

The first results concern chiral intermetallic compounds, promising materials for the field of enantioselective heterogeneous catalysis. Due to the intrinsic chirality, their surface can have a preference for a particular enantiomer of a chiral molecule. An example of such striking enantioselectivity with respect to the adsorption of 9-ethynylphenanthrene by the (111) surface of the PdGa chiral intermetallic compound has been demonstrated in recent experiments. Nevertheless, the theoretical explanation of this phenomenon has turned into a challenging task for modern modeling techniques. As it was shown in this work, the ability of chiral recognition of the material is a direct consequence of its chiral bulk structure. Since the surface layer of atoms is not chiral, to pin down the origin of selectivity we had to accurately estimate the dispersive interactions with the underlying layers. The failure of simplified DFT schemes based on atom-centered contributions was attributed to neglecting of the screening of vdW interactions within the bulk. The "surf" method, which corrects for that error, incorporates a screening effect through rescaling the parameters of the vdW potential. It would be important to test the approach for other interesting materials for the validation of the method and facilitating new discoveries in the field of enantioselective heterogeneous catalysis.

Interestingly, the $(\bar{1}\bar{1}\bar{1})$ surface of PdGa, located on the opposite side with respect

to (111), has a different structure, which induces a divergent mechanism of enantioselective recognition: the 9-ethynylphenanthrene molecules form trimers composed of units in the same enantiomeric form. Several ideas have been proposed in this work in order to explain the origin of trimer stabilization. Among them, only the hypothesis of the Ga adatom located in the center and keeping the monomers together turned out to be a feasible explanation, but this guess could only be accepted if the 9-ethynylphenanthrene was initially dehydrogenated. The hydrogenation state remains unclear so far, yet the formation of the molecular dimers on the surface suggests that, at an elevated temperature, 9-ethynylphenanthrene indeed loses the hydrogen of the ethynyl group.

Furthermore, the formation of domains by hexabenzocoronene on a Pt(111) surface was studied in this work. Experimentally it was observed that the molecule has two preferential adsorption configurations on the surface. This was investigated by means of DFT calculations and it was indeed found that hexabenzocoronene adsorbs in two possible symmetry-equivalent configurations. Experiments at high coverage revealed the repulsive nature of the intermolecular interactions, as the molecules tend to remain at far distances from each other. DFT simulations confirmed such a behavior in the form of an increase in the total energy of structures with closed-packed geometry. With the help of the charge density difference technique, it was demonstrated that the key phenomena that allows to explain such an observation is the electron transfer from the molecule to the topmost layer of the Pt. The overlap of the transfered charge-density causes Pauli repulsion, explaining the instability of the close-packed domains.

The next part of the thesis was devoted to the improvement of classical methods for nc-AFM simulations called ProbeParticle. The model allows to obtain reasonable nc-AFM images of adsorbed molecules in negligible time, although with several drawbacks. In the course of this work the model was ameliorated. A fitting procedure based on DFT calculations was developed and a general force field for the nc-AFM simulations was created. Usage of such a force field allowed to obtain simulated nc-AFM images in very good agreement with the ones produced by DFT and thus to the experimental results as well. However, as it was discussed in this

work, the model still has room for improvement. The created force field needs to be extended to include more commonly used chemical groups. The remaining discrepancies between the results of the ProbeParticle model and density functional calculations should be addressed either by introducing a more flexible potential or by introducing dynamic polarizability. The applicability of the developed approach should be tested for a larger variety of samples including the ones with the substrate. And finally, the possibility to apply our approach for the chemisorbed systems has to be investigated.

In the last part of the work, a new paradigm for a modern computational science is discussed. The large amount of data generated until today calls for conceptually new schemes of systematizing, sharing and validating the data. The unprecedented computational and storage resources available nowadays can be used to leverage the technologies such as big data analysis and machine learning for facilitating new discoveries. Unfortunately, the scientific data nowadays are quite scattered, therefore there is a need for systematization and providing a general way to access them. The last part of the thesis focuses on the AiiDA project, which by its design reflects the latest views on how modern computational science should be faced. It allows to automate routine operations, store the data in an efficient and systematic way, providing a possibility for easy sharing and machine learning. AiiDA is a very new, but quickly developing project, which is now at the core of the recently launched MaterialsCloud platform. The platform will allow to accelerate the discoveries of the new materials simplifying the storage and sharing of the research results and scientific know-hows in the form of workflows. Currently both AiiDA and the MaterialsCloud platform are open for research. They are evolving in a way that responds to the needs of the community adapting themselves to different use cases.

Bibliography

[1] Shinya Furukawa and Takayuki Komatsu. "Intermetallic Compounds: Promising Inorganic Materials for Well-Structured and Electronically Modified Reaction Environments for Efficient Catalysis". In: *ACS Catalysis* 7.1 (2017), pp. 735–765. DOI: 10.1021/acscatal.6b02603.

[2] Shinya Furukawa et al. "Selective Stereochemical Catalysis Controlled by Specific Atomic Arrangement of Ordered Alloys". In: *ChemCatChem* 7.21 (2015), pp. 3472–3479. DOI: 10.1002/cctc.201500808.

[3] Shinya Furukawa, Katsuya Takahashi, and Takayuki Komatsu. "Well-structured bimetallic surface capable of molecular recognition for chemoselective nitroarene hydrogenation". In: *Chemical Science* 7.7 (2016), pp. 4476–4484. DOI: 10.1039/C6SC00817H.

[4] Kirill Kovnir et al. "A new approach to well-defined, stable and site-isolated catalysts". In: *Science and Technology of Advanced Materials*. Chemical Physics of Solids 8.5 (2007), pp. 420–427. DOI: 10.1016/j.stam.2007.05.004.

[5] Jürgen Osswald et al. "Palladium–gallium intermetallic compounds for the selective hydrogenation of acetylene: Part I: Preparation and structural investigation under reaction conditions". In: *Journal of Catalysis* 258.1 (2008), pp. 210–218. DOI: 10.1016/j.jcat.2008.06.013.

[6] Jürgen Osswald et al. "Palladium–gallium intermetallic compounds for the selective hydrogenation of acetylene: Part II: Surface characterization and catalytic performance". In: *Journal of Catalysis* 258.1 (2008), pp. 219–227. DOI: 10.1016/j.jcat.2008.06.014.

[7] Marc Armbrüster et al. "Pd-Ga Intermetallic Compounds as Highly Selective Semihydrogenation Catalysts". In: *Journal of the American Chemical Society* 132.42 (2010), pp. 14745–14747. DOI: 10.1021/ja106568t.

[8] Christoph Rameshan et al. "In situ XPS study of methanol reforming on PdGa near-surface intermetallic phases". In: *Journal of Catalysis* 290 (2012), pp. 126–137. DOI: 10.1016/j.jcat.2012.03.009.

[9] Jan Prinz et al. "Highly Enantioselective Adsorption of Small Prochiral Molecules on a Chiral Intermetallic Compound". In: *Angewandte Chemie International Edition* 54.13 (2015), pp. 3902–3906. DOI: 10.1002/anie.201410107.

[10] Dirk Rosenthal et al. "Surface Investigation of Intermetallic PdGa(-1-1-1)". In: *Langmuir* 28.17 (2012), pp. 6848–6856. DOI: 10.1021/la2050509.

[11] John C. H. Spence et al. "On the minimum number of beams needed to distinguish enantiomorphs in X-ray and electron diffraction". In: *Acta Crystallographica Section A: Foundations of Crystallography* 50.5 (1994), pp. 647–650. DOI: 10.1107/S0108767394002850.

[12] Jan Prinz et al. "Isolated Pd Sites on the Intermetallic PdGa(111) and PdGa(-1-1-1) Model Catalyst Surfaces". In: *Angewandte Chemie* 124.37 (2012), pp. 9473–9477. DOI: 10.1002/ange.201203787.

[13] Peter Gille et al. "Growth of large PdGa single crystals from the melt". In: *Intermetallics* 18.8 (2010), pp. 1663–1668. DOI: 10.1016/j.intermet.2010.04.023.

[14] Jan Prinz et al. "Ensemble Effect Evidenced by CO Adsorption on the 3-Fold PdGa Surfaces". In: *The Journal of Physical Chemistry C* 118.23 (2014), pp. 12260–12265. DOI: 10.1021/jp501584f.

[15] Jan Prinz et al. "Adsorption of Small Hydrocarbons on the Three-Fold PdGa Surfaces: The Road to Selective Hydrogenation". In: *Journal of the American Chemical Society* 136.33 (2014), pp. 11792–11798. DOI: 10.1021/ja505936b.

[16] Mitch Jacoby. "2-D Stereoselectivity". In: *Chemical & Engineering News* 80.12 (2002), pp. 43–46.

[17] Andrew J. Gellman. "Chiral Surfaces: Accomplishments and Challenges". In: *ACS Nano* 4.1 (2010), pp. 5–10. DOI: 10.1021/nn901885n.

[18] Robert M. Hazen and David S. Sholl. "Chiral selection on inorganic crystalline surfaces". In: *Nature Materials* 2.6 (2003), pp. 367–374. DOI: 10.1038/nmat879.

[19] Yonghui Wang et al. "Enhancing Enantiomeric Separation with Strain: The Case of Serine on Cu(531)". In: *Journal of the American Chemical Society* 139.24 (2017), pp. 8167–8173. DOI: 10.1021/jacs.7b01216.

[20] Tugce Eralp et al. "The Importance of Attractive Three-Point Interaction in Enantioselective Surface Chemistry: Stereospecific Adsorption of Serine on the Intrinsically Chiral Cu{531} Surface". In: *Journal of the American Chemical Society* 134.23 (2012), pp. 9615–9621. DOI: 10.1021/ja210499m.

[21] Gerd Binnig and Heinrich Rohrer. "Scanning tunneling microscopy". In: *Surface Science* 126.1 (1983), pp. 236–244. DOI: 10.1016/0039-6028(83)90716-1.

[22] Gerd Binnig, Calvin F. Quate, and Christoph Gerber. "Atomic Force Microscope". In: *Physical Review Letters* 56.9 (1986), pp. 930–933. DOI: 10.1103/PhysRevLett.56.930.

[23] Seizo Morita et al., eds. *Noncontact Atomic Force Microscopy: Volume 3*. NanoScience and Technology. Springer International Publishing, 2015.

[24] Yves Martin, Clayton C. Williams, and H. Kumar Wickramasinghe. "Atomic force microscope – force mapping and profiling on a sub 100-Å scale". In: *Journal of Applied Physics* 61.10 (1987), pp. 4723–4729. DOI: 10.1063/1.338807.

[25] Dario Anselmetti et al. "Attractive-mode imaging of biological materials with dynamic force microscopy". In: *Nanotechnology* 5.2 (1994), p. 87. DOI: 10.1088/0957-4484/5/2/004.

[26] Franz J. Giessibl. "Atomic Resolution of the Silicon (111)-(7x7) Surface by Atomic Force Microscopy". In: *Science* 267.5194 (1995), pp. 68–71. DOI: 10.1126/science.267.5194.68.

[27] Yoshiaki Sugimoto et al. "Chemical identification of individual surface atoms by atomic force microscopy". In: *Nature* 446.7131 (2007), pp. 64–67. DOI: 10.1038/nature05530.

[28] Yoshiaki Sugimoto et al. "Complex Patterning by Vertical Interchange Atom Manipulation Using Atomic Force Microscopy". In: *Science* 322.5900 (2008), pp. 413–417. DOI: 10.1126/science.1160601.

[29] Leo Gross et al. "The Chemical Structure of a Molecule Resolved by Atomic Force Microscopy". In: *Science* 325.5944 (2009), pp. 1110–1114. DOI: 10.1126/science.1176210.

[30] Leo Gross et al. "Bond-Order Discrimination by Atomic Force Microscopy". In: *Science* 337.6100 (2012), pp. 1326–1329. DOI: 10.1126/science.1225621.

[31] Fabian Mohn et al. "Imaging the charge distribution within a single molecule". In: *Nature Nanotechnology* 7.4 (2012), pp. 227–231. DOI: 10.1038/nnano.2012.20.

[32] Prokop Hapala et al. "Mapping the electrostatic force field of single molecules from high-resolution scanning probe images". In: *Nature Communications* 7 (2016), p. 11560. DOI: 10.1038/ncomms11560.

[33] Prokop Hapala et al. "Mechanism of high-resolution STM/AFM imaging with functionalized tips". In: *Physical Review B* 90.8 (2014), p. 085421. DOI: 10.1103/PhysRevB.90.085421.

[34] Prokop Hapala et al. "Origin of High-Resolution IETS-STM Images of Organic Molecules with Functionalized Tips". In: *Physical Review Letters* 113.22 (2014), p. 226101. DOI: 10.1103/PhysRevLett.113.226101.

[35] Stefano Curtarolo et al. "AFLOWLIB.ORG: A distributed materials properties repository from high-throughput ab initio calculations". In: *Computational Materials Science* 58 (2012), pp. 227–235. DOI: 10.1016/j.commatsci.2012.02.002.

[36] *EUDAT - Research Data Services, Expertise & Technology Solutions*. URL: https://www.eudat.eu/.

[37] *NOMAD Repository*. URL: http://www.nomad-repository.eu/.

[38] Nikolai Zarkevich. "Structural database for reducing cost in materials design and complexity of multiscale computations." In: *Complexity* 11.4 (2006), pp. 36–42. DOI: 10.1002/cplx.20117.

[39] David D. Landis et al. "The Computational Materials Repository". In: *Computing in Science Engineering* 14.6 (2012), pp. 51–57. DOI: 10.1109/MCSE.2012.16.

[40] Pierre Villars et al. "The Pauling File, Binaries Edition". In: *Journal of Alloys and Compounds*. Proceedings of the VIII International Conference on Crystal Chemistry of Intermetallic Compounds 367.1 (2004), pp. 293–297. DOI: 10.1016/j.jallcom.2003.08.058.

[41] Sam Adams et al. "The Quixote project: Collaborative and Open Quantum Chemistry data management in the Internet age". In: *Journal of Cheminformatics* 3.1 (2011), p. 38. DOI: 10.1186/1758-2946-3-38.

[42] Anubhav Jain et al. "A high-throughput infrastructure for density functional theory calculations". In: *Computational Materials Science* 50.8 (2011), pp. 2295–2310. DOI: 10.1016/j.commatsci.2011.02.023.

[43] James E. Saal et al. "Materials Design and Discovery with High-Throughput Density Functional Theory: The Open Quantum Materials Database (OQMD)". In: *The Journal of The Minerals, Metals & Materials Society* 65.11 (2013), pp. 1501–1509. DOI: 10.1007/s11837-013-0755-4.

[44] Pedro R. C. da Silveira, Cesar R. S. da Silva, and Renata M. Wentzcovitch. "Metadata management for distributed first principles calculations in VLab–A collaborative cyberinfrastructure for materials computation". In: *Computer Physics Communications* 178.3 (2008), pp. 186–198. DOI: 10.1016/j.cpc.2007.09.001.

[45] Giovanni Pizzi et al. "AiiDA: automated interactive infrastructure and database for computational science". In: *Computational Materials Science* 111.Supplement C (2016), pp. 218–230. DOI: 10.1016/j.commatsci.2015.09.013.

[46] Giovanni Pizzi et al. *AiiDA: automated interactive infrastructure and database for computational science*. URL: http://www.aiida.net/.

[47] Pierre Hohenberg and Walter Kohn. "Inhomogeneous Electron Gas". In: *Physical Review* 136.3B (1964), B864–B871. DOI: 10.1103/PhysRev.136.B864.

[48] Walter Kohn and Lu J. Sham. "Self-Consistent Equations Including Exchange and Correlation Effects". In: *Physical Review* 140.4A (1965), A1133–A1138. DOI: 10.1103/PhysRev.140.A1133.

[49] Hans Hellman. *Einführung in die Quantenchemie*. Franz Deuticke, 1937.

[50] Richard P. Feynman. "Forces in Molecules". In: *Physical Review* 56.4 (1939), pp. 340–343. DOI: 10.1103/PhysRev.56.340.

[51] Max Born and J. Robert Oppenheimer. "Zur Quantentheorie der Molekeln". In: *Annalen der Physik* 389.20 (1927), pp. 457–484. DOI: 10.1002/andp.19273892002.

[52] Teepanis Chachiyo. "Communication: Simple and accurate uniform electron gas correlation energy for the full range of densities". In: *The Journal of Chemical Physics* 145.2 (2016), p. 021101. DOI: 10.1063/1.4958669.

[53] David M. Ceperley and Berni J. Alder. "Ground State of the Electron Gas by a Stochastic Method". In: *Physical Review Letters* 45.7 (1980), pp. 566–569. DOI: 10.1103/PhysRevLett.45.566.

[54] John P. Perdew, Kieron Burke, and Matthias Ernzerhof. "Generalized Gradient Approximation Made Simple". In: *Physical Review Letters* 77.18 (1996), pp. 3865–3868. DOI: 10.1103/PhysRevLett.77.3865.

[55] Axel D. Becke. "A new mixing of Hartree–Fock and local density-functional theories". In: *The Journal of Chemical Physics* 98.2 (1993), pp. 1372–1377. DOI: 10.1063/1.464304.

[56] Philip J. Stephens et al. "Ab Initio Calculation of Vibrational Absorption and Circular Dichroism Spectra Using Density Functional Force Fields". In: *The Journal of Physical Chemistry* 98.45 (1994), pp. 11623–11627. DOI: 10.1021/j100096a001.

[57] Carlo Adamo and Vincenzo Barone. "Toward reliable density functional methods without adjustable parameters: The PBE0 model". In: *The Journal of Chemical Physics* 110.13 (1999), pp. 6158–6170. DOI: 10.1063/1.478522.

[58] Matthias Ernzerhof and Gustavo E. Scuseria. "Assessment of the Perdew–Burke–Ernzerhof exchange-correlation functional". In: *The Journal of Chemical Physics* 110.11 (1999), pp. 5029–5036. DOI: 10.1063/1.478401.

[59] Joost VandeVondele, Urban Borštnik, and Jürg Hutter. "Linear Scaling Self-Consistent Field Calculations with Millions of Atoms in the Condensed Phase". In: *Journal of Chemical Theory and Computation* 8.10 (2012), pp. 3565–3573. DOI: 10.1021/ct200897x.

[60] Manuel Guidon et al. "Ab initio molecular dynamics using hybrid density functionals". In: *The Journal of Chemical Physics* 128.21 (2008), p. 214104. DOI: 10.1063/1.2931945.

[61] Manuel Guidon, Jürg Hutter, and Joost VandeVondele. "Auxiliary Density Matrix Methods for Hartree-Fock Exchange Calculations". In: *Journal of Chemical Theory and Computation* 6.8 (2010), pp. 2348–2364. DOI: 10.1021/ct1002225.

[62] Noa Marom et al. "Dispersion Interactions with Density-Functional Theory: Benchmarking Semiempirical and Interatomic Pairwise Corrected Density Functionals". In: *Journal of Chemical Theory and Computation* 7.12 (2011), pp. 3944–3951. DOI: 10.1021/ct2005616.

[63] Sándor Kristyán and Péter Pulay. "Can (semi)local density functional theory account for the London dispersion forces?" In: *Chemical Physics Letters* 229.3 (1994), pp. 175–180. DOI: 10.1016/0009-2614(94)01027-7.

[64] Henrik Rydberg et al. "Van der Waals Density Functional for Layered Structures". In: *Physical Review Letters* 91.12 (2003), p. 126402. DOI: 10.1103/PhysRevLett.91.126402.

[65] Vivekanand V. Gobre and Alexandre Tkatchenko. "Scaling laws for van der Waals interactions in nanostructured materials". In: *Nature Communications* 4 (2013), p. 2341. DOI: 10.1038/ncomms3341.

[66] Stefan Grimme. "Density functional theory with London dispersion corrections". In: *Wiley Interdisciplinary Reviews: Computational Molecular Science* 1.2 (2011), pp. 211–228. DOI: 10.1002/wcms.30.

[67] Stefan Grimme. "Accurate description of van der Waals complexes by density functional theory including empirical corrections". In: *Journal of Computational Chemistry* 25.12 (2004), pp. 1463–1473. DOI: 10.1002/jcc.20078.

[68] Stefan Grimme. "Semiempirical GGA-type density functional constructed with a long-range dispersion correction". In: *Journal of Computational Chemistry* 27.15 (2006), pp. 1787–1799. DOI: 10.1002/jcc.20495.

[69] Stefan Grimme et al. "A consistent and accurate ab initio parametrization of density functional dispersion correction (DFT-D) for the 94 elements H-Pu". In: *The Journal of Chemical Physics* 132.15 (2010), p. 154104. DOI: 10.1063/1.3382344.

[70] Qin Wu and Weitao Yang. "Empirical correction to density functional theory for van der Waals interactions". In: *The Journal of Chemical Physics* 116.2 (2001), pp. 515–524. DOI: 10.1063/1.1424928.

[71] Daniel J. Margoliash and William J. Meath. "Pseudospectral dipole oscillator strength distributions and some related two body interaction coefficients for H, He, Li, N, O, H2, N2, O2, NO, N2O, H2O, NH3, and CH4". In: *The Journal of Chemical Physics* 68.4 (1978), pp. 1426–1431. DOI: 10.1063/1.435963.

[72] Hendrik B. G. Casimir and Dirk Polder. "The Influence of Retardation on the London-van der Waals Forces". In: *Physical Review* 73.4 (1948), pp. 360–372. DOI: 10.1103/PhysRev.73.360.

[73] Anguang Hu, Darrin M. York, and Tom K. Woo. "Time-dependent density functional theory calculations of molecular static and dynamic polarizabilities, cauchy coefficients and their anisotropies with atomic numerical basis functions". In: *Journal of Molecular Structure: THEOCHEM* 591.1 (2002), pp. 255–266. DOI: 10.1016/S0166-1280(02)00246-4.

[74] Alexandre Tkatchenko and Matthias Scheffler. "Accurate Molecular Van Der Waals Interactions from Ground-State Electron Density and Free-Atom Reference Data". In: *Physical Review Letters* 102.7 (2009), p. 073005. DOI: 10.1103/PhysRevLett.102.073005.

[75] Xiang Chu and Alexander Dalgarno. "Linear response time-dependent density functional theory for van der Waals coefficients". In: *The Journal of Chemical Physics* 121.9 (2004), pp. 4083–4088. DOI: 10.1063/1.1779576.

[76] Alexandre Tkatchenko et al. "Accurate and Efficient Method for Many-Body van der Waals Interactions". In: *Physical Review Letters* 108.23 (2012), p. 236402. DOI: 10.1103/PhysRevLett.108.236402.

[77] Robert A. DiStasio, O. Anatole von Lilienfeld, and Alexandre Tkatchenko. "Collective many-body van der Waals interactions in molecular systems". In: *Proceedings of the National Academy of Sciences* 109.37 (2012), pp. 14791–14795. DOI: 10.1073/pnas.1208121109.

[78] Robert A. DiStasio, Vivekanand V. Gobre, and Alexandre Tkatchenko. "Many-body van der Waals interactions in molecules and condensed matter". In: *Journal of Physics: Condensed Matter* 26.21 (2014), p. 213202. DOI: 10.1088/0953-8984/26/21/213202.

[79] Alberto Ambrosetti et al. "Long-range correlation energy calculated from coupled atomic response functions". In: *The Journal of Chemical Physics* 140.18 (2014), 18A508. DOI: 10.1063/1.4865104.

[80] Alberto Ambrosetti et al. "Wavelike charge density fluctuations and van der Waals interactions at the nanoscale". In: *Science* 351.6278 (2016), pp. 1171–1176. DOI: 10.1126/science.aae0509.

[81] Martin A. Blood-Forsythe et al. "Analytical nuclear gradients for the range-separated many-body dispersion model of noncovalent interactions". In: *Chemical Science* 7.3 (2016), pp. 1712–1728. DOI: 10.1039/C5SC03234B.

[82] Victor G. Ruiz et al. "Density-Functional Theory with Screened van der Waals Interactions for the Modeling of Hybrid Inorganic-Organic Systems". In: *Physical Review Letters* 108.14 (2012), p. 146103. DOI: 10.1103/PhysRevLett.108.146103.

[83] Victor G. Ruiz, Wei Liu, and Alexandre Tkatchenko. "Density-functional theory with screened van der Waals interactions applied to atomic and molecular adsorbates on close-packed and non-close-packed surfaces". In: *Physical Review B* 93.3 (2016), p. 035118. DOI: 10.1103/PhysRevB.93.035118.

[84] Evgenii M. Lifshitz. "The Theory of Molecular Attractive Forces Between Solids". In: *Journal of Experimental and Theoretical Physics* 2 (1956), pp. 73–83.

[85] Eugene Zaremba and Walter Kohn. "Van der Waals interaction between an atom and a solid surface". In: *Physical Review B* 13.6 (1976), pp. 2270–2285. DOI: 10.1103/PhysRevB.13.2270.

[86] Sharad H. Patil, Kwong-Tin Tang, and J. Peter Toennies. "Damping functions for the pairwise sum model of the atom–surface potential". In: *The Journal of Chemical Physics* 116.18 (2002), pp. 8118–8123. DOI: 10.1063/1.1467050.

[87] Glenn J. Martyna and Mark E. Tuckerman. "A reciprocal space based method for treating long range interactions in ab initio and force-field-based calculations in clusters". In: *The Journal of Chemical Physics* 110.6 (1999), pp. 2810–2821. DOI: 10.1063/1.477923.

[88] Michele Parrinello et al. *CPMD*. URL: http://www.cpmd.org/.

[89] Paolo Giannozzi et al. "QUANTUM ESPRESSO: a modular and open-source software project for quantum simulations of materials". In: *Journal of Physics: Condensed Matter* 21.39 (2009), p. 395502. DOI: 10.1088/0953-8984/21/39/395502.

[90] Xavier Gonze et al. "ABINIT: First-principles approach to material and nanosystem properties". In: *Computer Physics Communications*. 40 YEARS OF CPC: A celebratory issue focused on quality software for high performance, grid and novel computing architectures 180.12 (2009), pp. 2582–2615. DOI: 10.1016/j.cpc.2009.07.007.

[91] José M. Soler et al. "The SIESTA method for ab initio order- N materials simulation". In: *Journal of Physics: Condensed Matter* 14.11 (2002), p. 2745. DOI: 10.1088/0953-8984/14/11/302.

[92] Volker Blum et al. "Ab initio molecular simulations with numeric atom-centered orbitals". In: *Computer Physics Communications* 180.11 (2009), pp. 2175–2196. DOI: 10.1016/j.cpc.2009.06.022.

[93] Andreas Marek et al. "The ELPA library: scalable parallel eigenvalue solutions for electronic structure theory and computational science". In: *Journal of Physics: Condensed Matter* 26.21 (2014), p. 213201. DOI: 10.1088/0953-8984/26/21/213201.

[94] Thomas Auckenthaler et al. "Parallel solution of partial symmetric eigenvalue problems from electronic structure calculations". In: *Parallel Computing*. 6th International Workshop on Parallel Matrix Algorithms and Applications (PMAA'10) 37.12 (2011), pp. 783–794. DOI: 10.1016/j.parco.2011.05.002.

[95] Ville Havu et al. "Efficient integration for all-electron electronic structure calculation using numeric basis functions". In: *Journal of Computational Physics* 228.22 (2009), pp. 8367–8379. DOI: 10.1016/j.jcp.2009.08.008.

[96] The CP2K developers group. *CP2K*. URL: http://www.cp2k.org/.

[97] Jorg Hutter et al. "Cp2k: Atomistic simulations of condensed matter systems". In: *Wiley Interdisciplinary Reviews: Computational Molecular Science* 4.1 (2014), pp. 15–25. DOI: 10.1002/wcms.1159.

[98] Joost VandeVondele et al. "Quickstep: Fast and accurate density functional calculations using a mixed Gaussian and plane waves approach". In: *Computer Physics Communications* 167.2 (2005), pp. 103–128. DOI: 10.1016/j.cpc.2004.12.014.

[99] Peter Schwerdtfeger. "The Pseudopotential Approximation in Electronic Structure Theory". In: *ChemPhysChem* 12.17 (2011), pp. 3143–3155. DOI: 10.1002/cphc.201100387.

[100] Charles G. Broyden. "The Convergence of a Class of Double-rank Minimization Algorithms 1. General Considerations". In: *IMA Journal of Applied Mathematics* 6.1 (1970), pp. 76–90. DOI: 10.1093/imamat/6.1.76.

[101] Roger Fletcher. "A new approach to variable metric algorithms". In: *The Computer Journal* 13.3 (1970), pp. 317–322. DOI: 10.1093/comjnl/13.3.317.

[102] Donald Goldfarb. "A family of variable-metric methods derived by variational means". In: *Mathematics of Computation* 24.109 (1970), pp. 23–26. DOI: 10.1090/S0025-5718-1970-0258249-6.

[103] David F. Shanno. "Conditioning of quasi-Newton methods for function minimization". In: *Mathematics of Computation* 24.111 (1970), pp. 647–656. DOI: 10.1090/S0025-5718-1970-0274029-X.

[104] Matthias Scheffler. "Thermodynamic Aspects of Bulk and Surface Defects–First-Principle Calculations". In: *Studies in Surface Science and Catalysis*. Ed. by J. Koukal. Vol. 40. Physics of Solid Surfaces 1987. Elsevier, 1988, pp. 115–122.

[105] Maurice Petyt. *Introduction to Finite Element Vibration Analysis*. Cambridge University Press, 2010.

[106] Oleg V. Gritsenko and Evert J. Baerends. "The analog of Koopmans' theorem for virtual Kohn–Sham orbital energies". In: *Canadian Journal of Chemistry* 87.10 (2009), pp. 1383–1391. DOI: 10.1139/V09-088.

[107] Delano P. Chong, Oleg V. Gritsenko, and Evert J. Baerends. "Interpretation of the Kohn–Sham orbital energies as approximate vertical ionization potentials". In: *The Journal of Chemical Physics* 116.5 (2002), pp. 1760–1772. DOI: 10.1063/1.1430255.

[108] Christoph Friedrich and Arno Schindlmayr. "Many-Body Perturbation Theory: The GW Approximation". In: *NIC Series* 31 (2006), p. 335.

[109] Frank A. Momany. "Determination of partial atomic charges from ab initio molecular electrostatic potentials. Application to formamide, methanol, and formic acid". In: *The Journal of Physical Chemistry* 82.5 (1978), pp. 592–601. DOI: 10.1021/j100494a019.

[110] Christopher I. Bayly et al. "A well-behaved electrostatic potential based method using charge restraints for deriving atomic charges: the RESP model". In: *The Journal of Physical Chemistry* 97.40 (1993), pp. 10269–10280. DOI: 10.1021/j100142a004.

[111] Daan Frenkel and Berend Smit. *Understanding Molecular Simulation*. Academic Press, 2002.

[112] Alessandro Laio and Francesco L. Gervasio. "Metadynamics: a method to simulate rare events and reconstruct the free energy in biophysics, chemistry and material science". In: *Reports on Progress in Physics* 71.12 (2008), p. 126601. DOI: 10.1088/0034-4885/71/12/126601.

[113] Alessandro Laio and Michele Parrinello. "Escaping free-energy minima". In: *Proceedings of the National Academy of Sciences* 99.20 (2002), pp. 12562–12566. DOI: 10.1073/pnas.202427399.

[114] Donald G. Truhlar, Bruce C. Garrett, and Stephen J. Klippenstein. "Current Status of Transition-State Theory". In: *The Journal of Physical Chemistry* 100.31 (1996), pp. 12771–12800. DOI: 10.1021/jp953748q.

[115] Eugene Wigner. "The transition state method". In: *Transactions of the Faraday Society* 34.0 (1938), pp. 29–41. DOI: 10.1039/TF9383400029.

[116] J. V. Barth et al. "Dynamics of Surface Migration in the Weak Corrugation Regime". In: *Physical Review Letters* 84.8 (2000), pp. 1732–1735. DOI: 10.1103/PhysRevLett.84.1732.

[117] Gregory Mills, Hannes Jónsson, and Gregory K. Schenter. "Reversible work transition state theory: application to dissociative adsorption of hydrogen". In: *Surface Science* 324.2 (1995), pp. 305–337. DOI: 10.1016/0039-6028(94)00731-4.

[118] Gregory Mills and Hannes Jónsson. "Quantum and thermal effects in H2 dissociative adsorption: Evaluation of free energy barriers in multidimensional quantum systems". In: *Physical Review Letters* 72.7 (1994), pp. 1124–1127. DOI: 10.1103/PhysRevLett.72.1124.

[119] Graeme Henkelman, Blas P. Uberuaga, and Hannes Jónsson. "A climbing image nudged elastic band method for finding saddle points and minimum energy paths". In: *The Journal of Chemical Physics* 113.22 (2000), pp. 9901–9904. DOI: 10.1063/1.1329672.

[120] Weinan E, Weiqing Ren, and Eric Vanden-Eijnden. "String method for the study of rare events". In: *Physical Review B* 66.5 (2002), p. 052301. DOI: 10.1103/PhysRevB.66.052301.

[121] Weinan E, Weiqing Ren, and Eric Vanden-Eijnden. "Simplified and improved string method for computing the minimum energy paths in barrier-crossing events". In: *The Journal of Chemical Physics* 126.16 (2007), p. 164103. DOI: 10.1063/1.2720838.

[122] John Bardeen. "Tunnelling from a Many-Particle Point of View". In: *Physical Review Letters* 6.2 (1961), pp. 57–59. DOI: 10.1103/PhysRevLett.6.57.

[123] J. Tersoff and D. R. Hamann. "Theory and Application for the Scanning Tunneling Microscope". In: *Physical Review Letters* 50.25 (1983), pp. 1998–2001. DOI: 10.1103/PhysRevLett.50.1998.

[124] C. Julian Chen. "Tunneling matrix elements in three-dimensional space: The derivative rule and the sum rule". In: *Physical Review B* 42.14 (1990), pp. 8841–8857. DOI: 10.1103/PhysRevB.42.8841.

[125] Jerry D. Tersoff. "Method for the calculation of scanning tunneling microscope images and spectra". In: *Physical Review B* 40.17 (1989), p. 11990.

[126] Roberto Gaspari et al. "s-orbital continuum model accounting for the tip shape in simulated scanning tunneling microscope images". In: *Physical Review B* 84.12 (2011), p. 125417. DOI: 10.1103/PhysRevB.84.125417.

[127] Leopold Talirz. *asetk*. URL: https://github.com/ltalirz/asetk.

[128] Nikolaj Moll et al. "The mechanisms underlying the enhanced resolution of atomic force microscopy with functionalized tips". In: *New Journal of Physics* 12.12 (2010), p. 125020. DOI: 10.1088/1367-2630/12/12/125020.

[129] Mark Neu et al. "Image correction for atomic force microscopy images with functionalized tips". In: *Physical Review B* 89.20 (2014), p. 205407. DOI: 10.1103/PhysRevB.89.205407.

[130] Mark P. Boneschanscher et al. “Sample Corrugation Affects the Apparent Bond Lengths in Atomic Force Microscopy”. In: *ACS Nano* 8.3 (2014), pp. 3006–3014. DOI: 10.1021/nn500317r.

[131] Chun-Sheng Guo et al. “High-Resolution Model for Noncontact Atomic Force Microscopy with a Flexible Molecule on the Tip Apex”. In: *The Journal of Physical Chemistry C* 119.3 (2015), pp. 1483–1488. DOI: 10.1021/jp511214e.

[132] Jun Zhang et al. “Real-Space Identification of Intermolecular Bonding with Atomic Force Microscopy”. In: *Science* 342.6158 (2013), pp. 611–614. DOI: 10.1126/science.1242603.

[133] Chun-Sheng Guo et al. “Origin of the Contrast Interpreted as Intermolecular and Intramolecular Bonds in Atomic Force Microscopy Images”. In: *The Journal of Physical Chemistry C* 119.25 (2015), pp. 14195–14200. DOI: 10.1021/acs.jpcc.5b02649.

[134] John E. Sader and Suzanne P. Jarvis. “Accurate formulas for interaction force and energy in frequency modulation force spectroscopy”. In: *Applied Physics Letters* 84.10 (2004), pp. 1801–1803. DOI: 10.1063/1.1667267.

[135] Michael Ellner et al. “The electric field of CO tips and its relevance for atomic force microscopy”. In: *Nano Letters* 16.3 (2016), pp. 1974–1980. DOI: 10.1021/acs.nanolett.5b05251.

[136] Nikolaj Moll et al. “Image Distortions of a Partially Fluorinated Hydrocarbon Molecule in Atomic Force Microscopy with Carbon Monoxide Terminated Tips”. In: *Nano Letters* 14.11 (2014), pp. 6127–6131. DOI: 10.1021/nl502113z.

[137] Aliaksandr V. Yakutovich et al. “Hidden Beneath the Surface: Origin of the Observed Enantioselective Adsorption on PdGa(111)”. In: *J. Am. Chem. Soc.* 140.4 (2017), pp. 1401–1408. DOI: 10.1021/jacs.7b10980.

[138] Wei Liu, Alexandre Tkatchenko, and Matthias Scheffler. “Modeling Adsorption and Reactions of Organic Molecules at Metal Surfaces”. In: *Accounts of Chemical Research* 47.11 (2014), pp. 3369–3377. DOI: 10.1021/ar500118y.

[139] Reinhard J. Maurer et al. “Adsorption structures and energetics of molecules on metal surfaces: Bridging experiment and theory”. In: *Progress in Surface Science* 91.2 (2016), pp. 72–100. DOI: 10.1016/j.progsurf.2016.05.001.

[140] Jing Kong et al. “Efficient computation of the dispersion interaction with density-functional theory”. In: *Physical Review A* 79.4 (2009), p. 042510. DOI: 10.1103/PhysRevA.79.042510.

[141] Oleg A. Vydrov and Troy Van Voorhis. “Nonlocal van der Waals density functional: The simpler the better”. In: *The Journal of Chemical Physics* 133.24 (2010), p. 244103. DOI: 10.1063/1.3521275.

[142] Jan Hermann, Robert A. DiStasio, and Alexandre Tkatchenko. “First-Principles Models for van der Waals Interactions in Molecules and Materials: Concepts, Theory, and Applications”. In: *Chemical Reviews* 117.6 (2017), pp. 4714–4758. DOI: 10.1021/acs.chemrev.6b00446.

[143] Jan Hermann, Dario Alfè, and Alexandre Tkatchenko. “Nanoscale pi–pi stacked molecules are bound by collective charge fluctuations”. In: *Nature Communications* 8 (2017), p. 14052. DOI: 10.1038/ncomms14052.

[144] Noa Marom et al. “Many-Body Dispersion Interactions in Molecular Crystal Polymorphism”. In: *Angewandte Chemie International Edition* 52.26 (2013), pp. 6629–6632. DOI: 10.1002/anie.201301938.

[145] Anthony M. Reilly and Alexandre Tkatchenko. "van der Waals dispersion interactions in molecular materials: beyond pairwise additivity". In: *Chemical Science* 6.6 (2015), pp. 3289–3301. DOI: 10.1039/C5SC00410A.

[146] Johannes Hoja, Anthony M. Reilly, and Alexandre Tkatchenko. "First-principles modeling of molecular crystals: structures and stabilities, temperature and pressure". In: *Wiley Interdisciplinary Reviews: Computational Molecular Science* 7.1 (2017), n/a–n/a. DOI: 10.1002/wcms.1294.

[147] Jess Wellendorff et al. "Density functionals for surface science: Exchange-correlation model development with Bayesian error estimation". In: *Physical Review B* 85.23 (2012), p. 235149. DOI: 10.1103/PhysRevB.85.235149.

[148] Yingkai Zhang and Weitao Yang. "Comment on "Generalized Gradient Approximation Made Simple"". In: *Physical Review Letters* 80.4 (1998), pp. 890–890. DOI: 10.1103/PhysRevLett.80.890.

[149] Jiří Klimeš, David R. Bowler, and Angelos Michaelides. "Chemical accuracy for the van der Waals density functional". In: *Journal of Physics: Condensed Matter* 22.2 (2010), p. 022201. DOI: 10.1088/0953-8984/22/2/022201.

[150] Riccardo Sabatini, Tommaso Gorni, and Stefano de Gironcoli. "Nonlocal van der Waals density functional made simple and efficient". In: *Physical Review B* 87.4 (2013), p. 041108. DOI: 10.1103/PhysRevB.87.041108.

[151] Giovanni Bussi, Alessandro Laio, and Michele Parrinello. "Equilibrium Free Energies from Nonequilibrium Metadynamics". In: *Physical Review Letters* 96.9 (2006), p. 090601. DOI: 10.1103/PhysRevLett.96.090601.

[152] Fabio Pietrucci. "Strategies for the exploration of free energy landscapes: Unity in diversity and challenges ahead". In: *Reviews in Physics* 2 (2017), pp. 32–45. DOI: 10.1016/j.revip.2017.05.001.

[153] Jeng-Da Chai and Martin Head-Gordon. "Long-range corrected hybrid density functionals with damped atom–atom dispersion corrections". In: *Physical Chemistry Chemical Physics* 10.44 (2008), pp. 6615–6620. DOI: 10.1039/B810189B.

[154] Pablo Bechthold et al. "The Electronic Structure and Bonding of Acetylene on PdGa(110)". In: *The Journal of Physical Chemistry C* 119.32 (2015), pp. 18229–18238. DOI: 10.1021/acs.jpcc.5b04214.

[155] Warren E. Pickett. "Pseudopotential methods in condensed matter applications". In: *Computer Physics Reports* 9.3 (1989), pp. 115–197. DOI: 10.1016/0167-7977(89)90002-6.

[156] Joost VandeVondele and Jürg Hutter. "Gaussian basis sets for accurate calculations on molecular systems in gas and condensed phases". In: *The Journal of Chemical Physics* 127.11 (2007), p. 114105. DOI: 10.1063/1.2770708.

[157] Stefan Goedecker, Michael Teter, and Jürg Hutter. "Separable dual-space Gaussian pseudopotentials". In: *Physical Review B* 54.3 (1996), pp. 1703–1710. DOI: 10.1103/PhysRevB.54.1703.

[158] Jerry D. Tersoff and Donald R. Hamann. "Theory of the scanning tunneling microscope". In: *Scanning Tunneling Microscopy*. Ed. by H. Neddermeyer. Perspectives in Condensed Matter Physics 6. Springer Netherlands, 1985, pp. 59–67.

[159] Claudia Ambrosch-Draxl and Jorge O. Sofo. "Linear optical properties of solids within the full-potential linearized augmented planewave method". In: *Computer Physics Communications* 175.1 (2006), pp. 1–14. DOI: 10.1016/j.cpc.2006.03.005.

[160] Guo-Xu Zhang et al. "van der Waals Interactions in Ionic and Semiconductor Solids". In: *Physical Review Letters* 107.24 (2011), p. 245501. DOI: 10.1103/PhysRevLett.107.245501.

[161] Nicholas Metropolis et al. "Equation of State Calculations by Fast Computing Machines". In: *The Journal of Chemical Physics* 21.6 (1953), pp. 1087–1092. DOI: 10.1063/1.1699114.

[162] Pascal Ruffieux et al. "Supramolecular columns of hexabenzocoronenes on copper and gold (111) surfaces". In: *Physical Review B* 66.7 (2002). DOI: 10.1103/PhysRevB.66.073409.

[163] Minjung Kim and James R. Chelikowsky. "CO tip functionalization in subatomic resolution atomic force microscopy". In: *Applied Physics Letters* 107.16 (2015), p. 163109. DOI: 10.1063/1.4934273.

[164] Dorothea Golze, Jürg Hutter, and Marcella Iannuzzi. "Wetting of water on hexagonal boron nitride@Rh(111): a QM/MM model based on atomic charges derived for nano-structured substrates". In: *Physical Chemistry Chemical Physics* 17.22 (2015), pp. 14307–14316. DOI: 10.1039/C4CP04638B.

[165] Melissa L. P. Price, Dennis Ostrovsky, and William L. Jorgensen. "Gas-phase and liquid-state properties of esters, nitriles, and nitro compounds with the OPLS-AA force field". In: *Journal of Computational Chemistry* 22.13 (2001), pp. 1340–1352. DOI: 10.1002/jcc.1092.

[166] William L. Jorgensen and Julian Tirado-Rives. "The OPLS [optimized potentials for liquid simulations] potential functions for proteins, energy minimizations for crystals of cyclic peptides and crambin". In: *Journal of the American Chemical Society* 110.6 (1988), pp. 1657–1666. DOI: 10.1021/ja00214a001.

[167] Nikolaj Moll et al. "A simple model of molecular imaging with noncontact atomic force microscopy". In: *New Journal of Physics* 14.8 (2012), p. 083023. DOI: 10.1088/1367-2630/14/8/083023.

[168] Kota Iwata et al. "Chemical structure imaging of a single molecule by atomic force microscopy at room temperature". In: *Nature Communications* 6 (2015), p. 7766. DOI: 10.1038/ncomms8766.

[169] Arieh Warshel, Mitsunori Kato, and Andrei V. Pisliakov. "Polarizable Force Fields: History, Test Cases, and Prospects". In: *Journal of Chemical Theory and Computation* 3.6 (2007), pp. 2034–2045. DOI: 10.1021/ct700127w.

[170] Matthias Krack and Michele Parrinello. "All-electron ab-initio molecular dynamics". In: *Physical Chemistry Chemical Physics* 2.10 (2000), pp. 2105–2112. DOI: 10.1039/B001167N.

[171] Gerald Lippert, Jürg Hutter, and Michele Parrinello. "The Gaussian and augmented-plane-wave density functional method for ab initio molecular dynamics simulations". In: *Theoretical Chemistry Accounts* 103.2 (1999), pp. 124–140. DOI: 10.1007/s002140050523.

[172] John A. Nelder and Roger Mead. "A Simplex Method for Function Minimization". In: *The Computer Journal* 7.4 (1965), pp. 308–313. DOI: 10.1093/comjnl/7.4.308.

[173] Margaret Wright. "Direct search methods: Once scorned, now respectable". In: *Numerical analysis*. Addison-Wesley, 1996, pp. 191–208.

[174] Eric Jones, Travis E. Oliphant, Pearu Peterson, et al. *SciPy: Open source scientific tools for Python*. 2001. URL: https://www.scipy.org/.

[175] Stéfan van der Walt, S. Chris Colbert, and Gaël Varoquaux. "The NumPy Array: A Structure for Efficient Numerical Computation". In: *Computing in Science & Engineering* 13.2 (2011), pp. 22–30. DOI: 10.1109/MCSE.2011.37.

[176] Travis E. Oliphant. "Python for Scientific Computing". In: *Computing in Science & Engineering* 9.3 (2007), pp. 10–20. DOI: 10.1109/MCSE.2007.58.

[177] K. Jarrod Millman and Michael Aivazis. "Python for Scientists and Engineers". In: *Computing in Science & Engineering* 13.2 (2011), pp. 9–12. DOI: 10.1109/MCSE.2011.36.

[178] John D. Hunter. "Matplotlib: A 2D Graphics Environment". In: *Computing in Science & Engineering* 9.3 (2007), pp. 90–95. DOI: 10.1109/MCSE.2007.55.

[179] Ole Schütt et al. *aiida-cp2k: CP2K plugin for the AiiDA provenance and workflow engine*. URL: https://github.com/cp2k/aiida-cp2k.

[180] *Materials Cloud*. URL: https://www.materialscloud.org/.

[181] *NCCR MARVEL*. URL: http://nccr-marvel.ch/.

[182] Ask Hjorth Larsen et al. "The atomic simulation environment—a Python library for working with atoms". In: *Journal of Physics: Condensed Matter* 29.27 (2017), p. 273002. DOI: 10.1088/1361-648X/aa680e.

[183] Marco Di Giovannantonio et al. "On-Surface Synthesis of Indenofluorene Polymers by Oxidative Five-Membered Ring Formation". In: *J. Am. Chem. Soc.* (2018). DOI: 10.1021/jacs.8b00587.

[184] Aliaksandr V. Yakutovich et al. "ProbeParticle: towards a fast and accurate model for nc-AFM simulations". In preparation.

[185] Marco Di Giovannantonio et al. "On-surface synthesis of open-shell Indenofluorenes: reactivity of radical sites". In preparation.

[186] Ole Schütt et al. "Jupyter@MaterialsCloud – an ecosystem to develop, execute and share scientific workflows". In preparation.

[187] Yves F. Forrer et al. "Growth and electronic structure of HBC on Pt(111) surface". In preparation.

[188] Samuel Stolz et al. "Formation and stability of 9-ethynylphenanthrene trimers on the PdGa:A(-1-1-1) surface". In preparation.

List of publications and author contributions

(I) Aliaksandr V. Yakutovich, Johannes Hoja, Daniele Passerone, Alexandre Tkatchenko, and Carlo A. Pignedoli. "Hidden Beneath the Surface: Origin of the Observed Enantioselective Adsorption on PdGa(111)". In: *J. Am. Chem. Soc.* 140.4 (Dec. 2017), pp. 1401–1408. ISSN: 0002-7863. DOI: 10.1021/jacs.7b10980.

The author performed the thermodynamics analysis, all abinitio geometry optimizations based on CP2K, STM and metadynamics simulations, analyzed the results and wrote the manuscript together with Carlo Pignedoli and Johannes Hoja.

(II) Marco Di Giovannantonio, José I. Urgel, Uliana Beser, Aliaksandr V. Yakutovich, Jan Wilhelm, Carlo A. Pignedoli, Pascal Ruffieux, Akimitsu Narita, Klaus Müllen, and Roman Fasel. "On-Surface Synthesis of Indenofluorene Polymers by Oxidative Five-Membered Ring Formation". In: *J. Am. Chem. Soc.* (Feb. 2018). ISSN: 0002-7863. DOI: 10.1021/jacs.8b00587.

The author performed nc-AFM simulations of indenofluorene polymers and assisted Marco Di Giovannantonio interpreting the nc-AFM experimental results.

Works in preparation

(III) Aliaksandr V. Yakutovich, Carlo A. Pignedoli, Prokop Hapala, Pavel Jelinek, Corina Keller, Marcella Iannuzzi, and Daniele Passerone. "ProbeParticle: towards a fast and accurate model for nc-AFM simulations". In preparation.

The author performed all DFT-AFM calculations, implemented the new "2-point-ProbeParticle" model, fitted parameters of both standard and new ProbeParticle to the DFT-AFM, analyzed the results and wrote the manuscript together with Prokop Hapala, Daniele Passerone, Pavel Jelinek and Carlo Pignedoli.

(IV) Marco Di Giovannantonio, Aliaksandr V. Yakutovich, José I. Urgel, Shantanu Mishra, Kristjan Eimre, Qiang Chen, Carlo A. Pignedoli, Pascal Ruffieux, Akimitsu Narita, Klaus Müllen, and Roman Fasel. "On-surface synthesis of open-shell Indenofluorenes: reactivity of radical sites". In preparation.

The author performed geometry optimizations, subsequent nc-AFM simulations and helped Marco Di Giovannantonio to identify the chemical composition of indenofluorene units.

(V) Ole Schütt, Edward Ditler, Aliaksandr V. Yakutovich, Daniele Passerone, and Pignedoli Carlo A. "Jupyter@MaterialsCloud – an ecosystem to develop, execute and share scientific workflows". In preparation.

The author implemented the initial version of the workflow to study electronic properties of nanoribbons and contributed to the development of the CP2K plugin.

(VI) Yves F. Forrer, Aliaksandr V. Yakutovich, Thomas Dienel, Roland Widmer, Liwei Liu, Daniele Passerone, Carlo A. Pignedoli, and Oliver Gröning. "Growth and electronic structure of HBC on Pt(111) surface". In preparation.

The author performed DFT calculations, interpreted the results, helped to identify the adsorption sites of Hexabenzocoronene molecules on Pt(111) surface and to explain the instability of the close-packed molecular domains.

(VII) Samuel Stolz, Aliaksandr V. Yakutovich, Roland Widmer, Daniele Passerone, Carlo A. Pignedoli, and Oliver Gröning. "Formation and stability of 9-ethynylphenanthrene trimers on the PdGa:A(-1-1-1) surface". In preparation.

The author performed all DFT calculations, DFT-based STM simulations and nc-AFM simulations to explain the stability of the 9-EP trimers on PdGa:A(-1-1-1) surface.

Acknowledgements

My "slow train to Switzerland"[4] started at the end of June 2013. At that time I was doing my internship at the Sanofi pharmaceutical company located in Montpellier – a beautiful southern city of France. After a short exchange of emails followed by two skype-interviews, I was informed by Daniele Passerone that I am accepted for a PhD position at the nanotech@surafces laboratory at Empa.

The starting project, which turned out to be quite a challenging task and lasted basically my whole stay at Empa, was to understand the enantioselectivity of the (111) surface of a PdGa intermetallic compound. It was a journey full of mistakes, moments of excitements about "almost finding an explanation" followed by moments of great discouragement, when every "definitely right" hypothesis ended up to be wrong. The question was rather simple: we had to understand why a prochiral molecule (9-ehtynylphenanthrene) adsorbs enantioselectively on the (111) surface of PdGa.

When things do not work as expected, one starts to question everything: the atomic model, the method, its accuracy, the validity of adopted approximations, etc...We were looking at the certain (minimal) value of the energy difference, but all the attempts did not provide enough. Trying to correct ourselves for all possible mistakes, we arrived at the only possible explanation: the enantioselectivity originates from the vdW interaction. But it was still just a guess, and the best tool available at that time for such a large system as ours was the random phase approximation. However, we were not sure whether this approach works for metals, and the only way to check was simply to try to use it. Such a try meant for us that even for the smallest reasonable size of the system we would need to reserve for one day one of the largest supercomputers in the world. In addition, those two calculations would completely use our quota for 3 months. I did this on a Sunday, in order to annoy as few people as possible. The result was discouraging again: the trimester quota was used up, the method turned out not to be working for metals. Quite a sad and discouraging situation.

[4]Inspired by the book of Diccon Bewes: "Slow Train to SWITZERLAND".

I am providing so many details because in order to appreciate the good people who surrounded me during my PdD, I wanted to describe how hard it was sometimes. It is thanks to Carlo Pignedoli and Daniele Passerone, who supervised my work the whole way through, that the problem finally got solved. I appreciate their availability for discussions, support in trying out new ideas and, most importantly, for not letting me to give up on the topic. It was their idea (after all those resultless attempts) to send me to Berlin to the group of Alexandre Tkatchenko – an expert of modeling the vdW interactions. Even though I did not meet him personally – I would like to thank him for the great pieces of advice concerning the methodology. I would also like to thank Johannes Hoja, who implemented the vdW$^{\text{surf}}$ method that finally provided a high enough energy difference to explain the experimentally observed enantioselectivity.

Furthermore, I would like to thank people from our experimental section. First of all I want to thank Jan Prinz. This work would not be possible without his reliable experiments with such a striking outcome – if we were not so convinced by his experimental results, we would not question our theoretical methods. I want to thank Roland Widmer, under whose supervision all the PdGa experiments were done. Another project related to the $(\bar{1}\bar{1}\bar{1})$ surface of PdGa was to understand why the very same 9-ethynylphenanthrene molecule in that case forms trimers. Unfortunately, we were not really able to solve this problem even though the amount of work and thoughts put into this was very large. Basically, by no available means of simulations and experiment could we find an answer to why the molecular trimers are kept stable on that surface. I would like to thank Samuel Stolz who did most of the experiments related to this surface. I really appreciate his ability to work hard and efficiently. I want to thank Oliver Gröening, who gave me a lot of suggestions regarding interpretation of experimental results and who was always available for discussions about the thermodynamics of processes at surfaces. The experiments of HBC adsorption on a Pt(111) surface was done by Yves Forrer. I want to thank him for the careful analysis that he had performed. I want to thank Yves for his hard work and very interesting results he had obtained.

I also interacted with many people for the nc-AFM simulatons. First of all I want

to thank Prokop Hapala, who developed the initial version of the ProbeParticle code and his supervisor, Pavel Jelinek, who was kind enough to host me in his laboratory in Prague for a one-week period. My thanks go to Corina Keller and Marcella Iannuzzi who implemented the DFT-based nc-AFM model, which I extensively used in my research. I also appreciate the collaboration with Marco Di Giovannantonio, a great person and scientist, who extensively used my implementation of the ProbeParticle model.

My personal thanks to Giovanni Pizzi – a main developer of the AiiDA code, who was able to convince me of the usefulness of AiiDA and who helped me to quickly start using it. I thank Leonid Kahle, who helped in creating the initial version of the CP2K-AiiDA plugin. I appreciate the help of Andrea Cepellotti and Nicolas Mounet in writing the workflow to study the electronic properties of 1D materials. I also thank Ole Shütt who took over the development of both projects, reshaped and considerably improved them.

I further thank Leopold Talirz, who helped me to start using his STM and asetk codes. I thank him and Prashant Shinde for sharing a great time in the same office and for their availability to help and to discuss scientific problems.

I kindly thank Jürg Hutter, who agreed to be my doctoral adviser. I thank the whole nanotech@surface group and the the head of the laboratory, Roman Fasel, for hosting me and for a collaborative and encouraging working atmosphere.

Finally, I would like to thank the Swiss National Science foundation and NCCR-MARVEL for the financial support. I would like to thank Swiss National Supercomputing Center for providing a powerful and reliable computation resources.


CMSZH
CHEMICAL MOLECULAR SCIENCES ZURICH
GRADUATE SCHOOL